(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
3 May 2001 (03.05.2001)

PCT

(10) International Publication Number
WO 01/31342 A2

(51) International Patent Classification[7]: G01N 33/574, C12Q 1/68, A61P 35/00

(21) International Application Number: PCT/US00/29556

(22) International Filing Date: 26 October 2000 (26.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/161,687	26 October 1999 (26.10.1999)	US
60/167,556	23 November 1999 (23.11.1999)	US
09/696,733	25 October 2000 (25.10.2000)	US

(71) Applicant: **TULARIK INC.** [US/US]; Two Corporate Drive, South San Fransisco, CA 94080 (US).

(72) Inventors: **POWERS, Scott**; 6 Westbrook Court, Greenlawn, NY 11740 (US). **SERVOSS, Allyson**; 15 Erin Lane, Setauket, NY 11733 (US). **ZENG, Rong, Wendy**; 70 Dix Highway, Dix Hills, NY 11746 (US).

(74) Agents: **KEZER, William, B.** et al.; Townsend and Townsend and Crew LLP, Two Embarcadero Center, 8th floor, San Francisco, CA 94111-3834 (US).

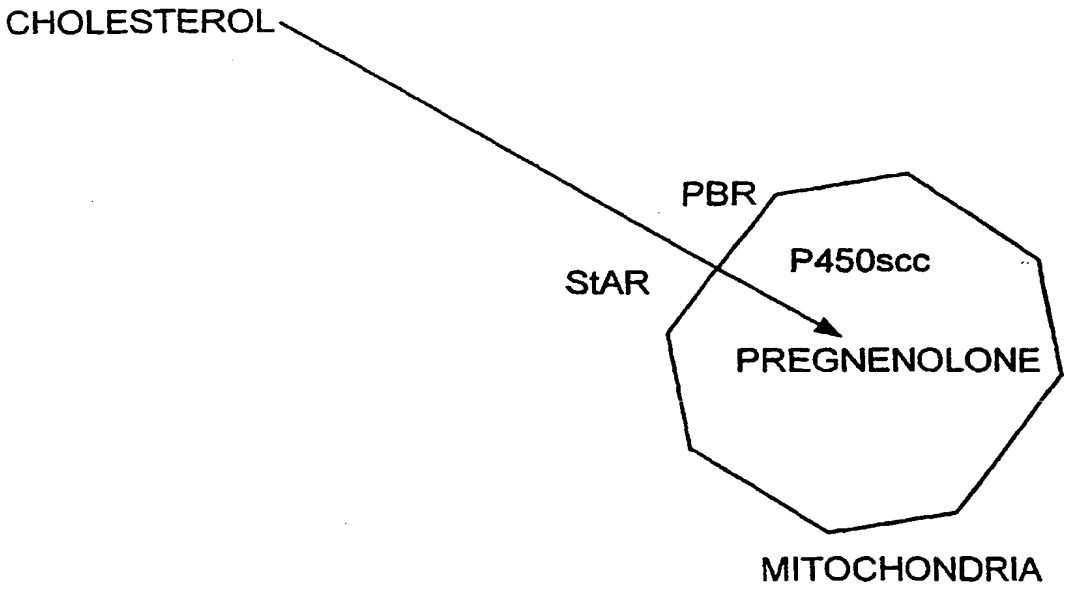
(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: DIAGNOSIS AND TREATMENT OF CANCER USING StAR POLYPEPTIDES AND POLYNUCLEOTIDES



(57) Abstract: The present invention provides methods, reagents, and kits for diagnosing and treating cancer in a mammal, *e.g.*, a human. This invention is based upon the surprising discovery that the StAR gene is overexpressed and/or amplified in a number of types of cancer cells. Accordingly, the present methods can be used to detect cancer or a propensity to develop cancer, to monitor the efficacy of a cancer treatment, and to treat cancer, *e.g.*, by inhibiting the expression and/or activity of StAR in a cancer cell.

WO 01/31342 A2

DIAGNOSIS AND TREATMENT OF CANCER USING StAR POLYPEPTIDES AND POLYNUCLEOTIDES

BACKGROUND OF THE INVENTION

Field Of The Invention

This invention relates to the field of cancer diagnosis and treatment. Methods and diagnostic reagents are provided for diagnosing and treating cancers that involve abnormal or abnormally expressed StAR nucleic acids or polypeptides.

Background

Despite years of research into its causes and potential treatments, cancer remains the second leading cause of death in the United States. Currently, more than 500,000 Americans die of cancer each year, more than 1,500 people a day. In addition, even more cases will be detected; for instance, in 1999 more than 1.2 million new cancer cases will be diagnosed. Of the many types of cancer, epithelial cancers are among the most prevalent and deadly. For example, about 175,000 new invasive cases of breast cancer are expected to occur among women in the United States during 1999, and more than 43,000 women will die of this disease. Although some progress has been made towards understanding the causes of various types of cancer, a major need remains for new tools for the diagnosis and treatment of this disease.

Cancer is a genetic disease of single cell origin caused by the accumulation of inherited and acquired mutations in specific cancer genes, which have normal cellular functions, but when mutated or present at abnormally high levels contribute to cancer. One type of mutation that unlocks the cancer-causing potential of a cancer gene is gene amplification or overexpression, *i.e.*, where a specific chromosomal region (including the cancer gene) has undergone a relative increase in DNA copy number, such that more copies of the cancer gene are present, or where the level of expression of a gene is increased, such that a correspondingly higher amount of mRNA and protein is produced, causing deleterious effects. Gene amplification is one of the primary mutational mechanisms for causing

primary genetic alterations in solid tumors, and most of the chromosomal regions that undergo amplification are not well characterized and do not harbor known oncogenes (Knuutila *et al.*, (1998) *Am. J. Pathol.*, 152:1107-23; Knuutila *et al.*, (1998) *Cancer Genet. Cytogenet.*, 100:25-30). Discovery of these amplified cancer genes will provide novel targets for diagnostic and therapeutic applications. In breast cancers, over 20 high-level amplified regions have been identified by CGH, but only three have been firmly associated with established oncogenes (HER2 at 17q12-q21, MYC at 8q24, and BCL1/Cyclin D1 at 11q13) (Knuutila *et al.*, (1998) *Am. J. Pathol.*, 152:1107-23). Accordingly, a large number of genes that are amplified and/or overexpressed in cancer cells remain to be identified.

One particularly effective method of identifying amplified regions of the genome is through representational difference analysis, or RDA (Lisitsyn *et al.*, (1993) *Science*, 259:946-951; U.S. Patent Nos. 5,436,142 and 5,501,964). Briefly, RDA involves several steps, including isolating a "representative" population of nucleic acid from a cell or cell type or interest, *e.g.*, a cancer cell. This "representative" population contains significantly less complexity than that found in the overall genome. Using a simultaneous hybridization and amplification step to enrich for those sequences not present in a control, or "tester" cell or cell population, RDA allows for the identification of sequences that are over-represented in the cell type of interest.

The steroidogenic acute regulatory protein, or "StAR," was identified in 1994 (Clark *et al.*, (1994) *J. Biol. Chem.*, 269:28314-22) based on its hormone-induced expression as well as its localization to the mitochondria. StAR is the rate-limiting protein in steroidogenesis, allowing entry of cholesterol into the mitochondria for the first enzymatic steps carried out by p450scc for hormone biosynthesis. The expression of StAR is normally strictly regulated, and usually only occurs in the presence of hormones, *e.g.*, trophic hormones.

The present invention is based on the surprising discovery that StAR polynucleotide sequences are amplified and/or overexpressed in many types of cancer cells in mammals. As described herein, this discovery has provided novel and badly needed diagnostic, prognostic, and therapeutic tools for many types of cancers.

SUMMARY OF THE INVENTION

The present invention provides methods, reagents, and kits for diagnosing and treating cancer in a mammal, *e.g.*, a human. This invention is based upon the surprising

discovery that the gene encoding the steroidogenic acute regulatory protein, or "StAR," is overexpressed and/or amplified in a number of types of cancer cells. Accordingly, the present methods can be used to detect cancer or a propensity to develop cancer, to monitor the efficacy of a cancer treatment, and to treat cancer, *e.g.*, by inhibiting the expression and/or activity of StAR in a cancer cell.

In one aspect, the present invention provides a method of detecting cancer in a biological sample from a mammal, the method comprising detecting the presence or absence of a StAR polypeptide or polynucleotide in the biological sample, wherein a diagnostic presence of the StAR polypeptide or polynucleotide indicates the presence of cancer in the biological sample.

In another aspect, the present invention provides a method of detecting cancer in a biological sample from a mammal, the method comprising detecting the level of a StAR polypeptide or polynucleotide in the biological sample, wherein an increased level of the StAR polypeptide or polynucleotide in the biological sample compared to the level expected in a control sample from a normal, cancer-free mammal indicates the presence of cancer in the biological sample.

In one embodiment, the detection is performed under conditions that would not detect the StAR polypeptide or polynucleotide in a sample that is not cancerous, and wherein a diagnostic presence or an increased level comprises detecting any StAR polypeptide or polynucleotide in the biological sample.

In another embodiment, the diagnostic presence or increased level comprises at least a 2-fold increase in the level of the StAR polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free mammal. In another embodiment, the diagnostic presence or increased level comprises at least a 5-fold increase in the level of the StAR polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free mammal. In another embodiment, the diagnostic presence or increased level comprises at least a 10-fold increase in the level of the StAR polypeptide or polynucleotide in the biological sample compared to a level expected in a sample from a control, cancer-free mammal.

In another embodiment, the method comprises detecting a StAR polypeptide by immunoassay. In another embodiment, the immunoassay comprises

immunohistochemistry. In another embodiment, the method comprises detecting the presence or absence of StAR gene amplification. In another embodiment, the method comprises detecting a StAR mRNA.

In another embodiment, the cancer is an epithelial cancer. In another embodiment, the epithelial cancer is selected from the group consisting of breast cancer, lung cancer, colorectal cancer, prostate cancer, kidney cancer, stomach cancer, bladder cancer, ovarian cancer, and cancer of the gastrointestinal tract. In another embodiment, the mammal is a human. In another embodiment, the biological sample is selected from the group consisting of tissue biopsy, blood sample, and nipple discharge.

In another aspect, the present invention provides a method of monitoring the efficacy of a cancer treatment, the method comprising detecting the level of a StAR polypeptide or polynucleotide in a biological sample from a mammal undergoing treatment for cancer, wherein a reduced level of the StAR polypeptide or polynucleotide in the biological sample compared to a level in a sample taken from the mammal prior to, or earlier in, the treatment is indicative of efficacious treatment.

In one embodiment, the cancer is an epithelial cancer. In another embodiment, the epithelial cancer is a breast cancer. In another embodiment, the mammal is a human.

In another aspect, the present invention provides methods for decreasing the proliferation of a cell with an elevated level of StAR, the method comprising reducing the level of StAR activity in the cell using an inhibitor of StAR.

In one embodiment, the cell is a cancer cell. In another embodiment, the cancer cell is an epithelial cancer cell. In another embodiment, the epithelial cancer cell is a breast cancer cell. In another embodiment, the inhibitor produces a decrease in the production of steroid hormones by the cell.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a physical map of the 8p11-p12 amplified chromosomal region, indicated by the solid line marked with solid circles. Underneath are solid lines representing individual YACs, and a dashed line representing genomic sequence from GenBank (Accession numbers: AP000065 through AP000084). STSs on this map (solid circles) were derived from RDA probes or from the genomic sequence or public databases at the Whitehead Institute, Stanford University, and the Sanger Center. The STSs are ordered

on the basis of genomic sequence; or by determination by PCR of their presence on these YACs and others not shown, and confirmed by radiation hybrid mapping using the GB3 panel (Stanford University RH server).

Figure 2 provides a schematic illustration of the role of StAR in steroid hormone biosynthesis.

DETAILED DESCRIPTION OF THE INVENTION AND PREFERRED EMBODIMENTS

I. Introduction

The present invention provides methods of diagnosing and treating cancer in mammals and other animals. These methods are based upon the discovery that the StAR gene is amplified and/or overexpressed in cancer cells. Accordingly, cancer cells can be detected by virtue of amplification or overexpression of StAR. In addition, the present invention provides methods for treating cancer, for example, by reducing the level of expression or activity of StAR in a cancer cell.

The methods of the invention typically involve detecting the presence or absence of StAR in a biological sample taken from a mammal. In numerous embodiments, the level of StAR polypeptide or polynucleotide is detected in a sample and compared to a "control" level, where a "control" level is a level of StAR that is actually detected in a cancer-free sample or a level that is expected for a typical cancer-free sample. Such detection may involve quantitative or qualitative detection of the polypeptide or polynucleotide, and may involve an actual comparison with a control value or, alternatively, may be performed so that the detection itself inherently indicates an increased level, or "diagnostic presence" of the StAR. As used herein, a "diagnostic presence" indicates any level of StAR that is greater than that expected in a noncancerous sample. In a presently preferred embodiment, assays for a StAR polypeptide or polynucleotide in a biological sample are conducted under conditions wherein a normal level of StAR polypeptide or polynucleotide, *i.e.*, a level typical of a noncancerous sample, *i.e.*, cancer-free, would not be detected. In such assays, therefore, the detection of any StAR polypeptide or nucleic acid in the biological sample indicates a diagnostic presence, or increased level. Typically, a diagnostic presence of StAR polypeptide or nucleic acid represents at least about a 1.5, 2, 5, 10, or greater fold increase in the StAR polypeptide or polynucleotide in the biological

sample compared to a level expected in a noncancerous sample. Detection of StAR can be performed *in vitro, i.e.*, in cells within a biological sample taken from the mammal, or *in vivo*.

According to the present methods, an increased level, or diagnostic presence, of StAR in a biological sample can be detected using any of a variety of forms of StAR, for example, StAR protein, StAR mRNA, or StAR genomic DNA, and can indicate an increase in, *e.g.*, StAR gene copy number, StAR transcription, StAR translation, StAR RNA or protein stability, or StAR protein activity. Accordingly, any of a large number of assays can be used to detect the StAR nucleic acids or polypeptides, such as StAR nucleic acid probes to detect StAR polynucleotides, or StAR-specific antibodies or other specific binding reagents to detect StAR polypeptides.

The ability to detect cancer cells by virtue of an increased level, or "diagnostic presence," of StAR is useful for any of a large number of applications. For example, an increased level or diagnostic presence of StAR in cells of a mammal can be used, alone or in combination with other diagnostic methods, to diagnose cancer in the mammal or to determine the propensity of a mammal to develop cancer over time. The detection of StAR can also be used to monitor the efficacy of a cancer treatment. For example, a level of a StAR polypeptide or polynucleotide after an anti-cancer treatment is compared to the level in the mammal before the treatment, wherein a decrease in the level of the StAR polypeptide or polynucleotide after the treatment indicates efficacious treatment.

An increased level or diagnostic presence of StAR can also be used to influence the choice of anti-cancer treatment in a mammal, where, for example, a strong increase in a level of StAR indicates the use of a more aggressive anti-cancer therapy, and a small increase or no increase indicates the use of a less aggressive anti-cancer therapy. In preferred embodiments, a level of StAR is used to guide the choice of an anti-cancer agent, in particular agents that affect the level or activity of a steroid hormone. For example, an increased level of StAR in breast or prostate tumor cells can indicate that the use of an agent that lowers the level or activity of estrogen or testosterone, respectively, would be effective in treating the tumor. In a preferred embodiment, an increased or diagnostic level of StAR as detected by immunoassay in a breast tumor is used to indicate the use of tamoxifen in the treatment of the tumor.

In addition, the ability to detect cancer cells can be useful to monitor the number or location of cancer cells in a patient, *in vivo* or *in vitro*, for example, to monitor the progression of the cancer over time. In addition, the level or presence or absence of StAR can be statistically correlated with the efficacy of particular anti-cancer therapies or with observed prognostic outcomes, thereby allowing the development of databases based on which a statistically-based prognosis, or a selection of the most efficacious treatment, can be made in view of a particular level or diagnostic presence of StAR.

The present invention also provides methods for treating cancer. In certain embodiments, the growth, proliferation, or steroid production of a cell with an elevated level of StAR polynucleotides, polypeptides, or polypeptide activity is inhibited. The growth, proliferation, or steroid production is decreased by, for example, contacting the cell with an inhibitor of StAR transcription or translation, or an inhibitor of the activity of a StAR polypeptide. Such inhibitors include, but are not limited to, antisense polynucleotides, ribozymes, antibodies, dominant negative StAR polypeptides, and small molecule inhibitors of StAR activity.

The present methods can be used to diagnose, determine the prognosis for, or treat, any of a number of types of cancers. In preferred embodiments, the cancer is an epithelial cancer, *e.g.*, breast, lung, colorectal, prostate, kidney, stomach, bladder, or ovarian cancer, or any cancer of the gastrointestinal tract.

The diagnostic methods of this invention can be used in animals including, for example, primates, canines, felines, murines, bovines, equines, ovines, porcines, lagomorphs, *etc*, as well as in humans. Suitable biological samples include, but are not limited to, tissue biopsy, blood sample, buccal scrape, saliva, nipple discharge, and urine.

Kits are also provided for carrying out the herein-disclosed diagnostic and therapeutic methods.

II. Definitions

The phrase "detecting a cancer" refers to the ascertainment of the presence or absence of cancer in an animal. "Detecting a cancer" can also refer to obtaining indirect evidence regarding the likelihood of the presence of cancerous cells in the animal. Detecting a cancer can be accomplished using the methods of this invention alone, in combination with other methods, or in light of other information regarding the state of health of the animal.

A "cancer" in an animal refers to the presence of cells possessing characteristics typical of cancer-causing cells, such as uncontrolled proliferation, immortality, metastatic potential, rapid growth and proliferation rate, and certain characteristic morphological features. Often, cancer cells will be in the form of a tumor, but such cells may exist alone within an animal, or may circulate in the blood stream as independent cells, such as leukemic cells.

"Providing a biological sample" means to obtain a biological sample for use in the methods described in this invention. Most often, this will be done by removing a sample of cells from an animal, but can also be accomplished by using previously isolated cells (*e.g.*, isolated by another person, at another time, and/or for another purpose), or by performing the methods of the invention *in vivo*.

A "control sample" refers to a sample of biological material representative of healthy, cancer-free animals. The level of StAR in a control sample is desirably typical of the general population of normal, cancer-free animals. This sample can be removed from an animal expressly for use in the methods described in this invention, or can be any biological material representative of normal, cancer-free animals. A control sample can also refer to an established level of StAR, representative of the cancer-free population, that has been previously established based on measurements from normal, cancer-free animals. If a detection method is used that only detects StAR when a level higher than that typical of a normal, cancer-free animal is present, *i.e.*, an immunohistochemical assay giving a simple positive or negative result, this is considered to be assessing the StAR level in comparison to the control level, as the control level is inherent in the assay.

A level of StAR polypeptide or polynucleotide that is "expected" in a control sample refers to a level that represents a typical, cancer-free sample, and from which an elevated, or diagnostic, presence of StAR polypeptide or polynucleotide can be distinguished. Preferably, an "expected" level will be controlled for such factors as the age, sex, medical history, *etc.* of the mammal, as well as for the particular biological sample being tested.

An "increased," or "elevated," level of StAR refers to a level of StAR polynucleotide or polypeptide, that, in comparison with a control level of StAR, is detectably higher. The method of comparison can be statistical, using quantified values for the level of

StAR, or can be compared using nonstatistical means, such as by a visual, subjective assessment by a human.

A "StAR polynucleotide" or "StAR nucleic acid" is a DNA or RNA sequence of at least about 50 nucleotides that is at least about 70% identical, preferably at least about 80%, or more, identical over a region of at least about 50, 100, 200, 500, or more nucleotides to a StAR polynucleotide sequence, from a mammal including, but not limited to, human, rat, mouse, hamster, cow, pig, horse, sheep, or any mammal, as, *e.g.*, provided in Clark *et al.*, (1994) *J. Biol. Chem.*, 269:28314-22; Mizutani *et al.*, (1997) *Life Sci.*, 61:1497-506; Rust *et al.*, (1998) *J. Mol. Endocrinol.*, 21(2):189-200; Fleury *et al.*, (1998) *J. Mol. Endocrinol.*, 21(2):131-9; Pilon *et al.*, (1997) *Endocrinology*, 138:1085-91; Lee *et al.*, (1997) *Biochem. Biophys. Res. Commun.*, 230(3):528-32; Hartung *et al.*, (1995) *Biochem. Biophys. Res. Commun.*, 215(2):646-53; and, *e.g.*, in GenBank Accession Nos: U66490, U53020, AB001349, L36062, L36061, S73839, AF031696, AF086814, Y17260, AB006007, AF038553, and NM_000349, or derivatives or fragments therof, or which specifically hybridizes to any of these polynucleotides. Typically, a "StAR polynucleotide" encodes a "StAR polypeptide." In preferred embodiments, a human StAR polynucleotide will comprise at least about 50%, 70%, 80%, 90%, or more nucleotide sequence identity to SEQ ID NO:1. Typically, the human StAR polynucleotide will comprise the endogenous StAR gene, present at genomic location 8p11.2, or will comprise an RNA transcribed from the endogenous StAR gene. A "StAR polynucleotide" can comprise naturally occurring nucleotides, or any derivative or analog thereof, *e.g.*, labeled or modified deoxyribo- or ribonucleotides. The term "StAR polynucleotide" can refer to a mutated copy of any of the above sequences, or a fragment thereof, and can refer to any naturally occurring or synthetic StAR sequence.

A "StAR protein" or "StAR polypeptide" refers to a polypeptide of at least about 20 amino acids that is typically about 70% identical, preferably at least about 80%, more preferably at least about 90%, or more, identical over a region of at least about 20, 50, 100 or more amino acids, to one or more StAR polypeptide sequences from any mammal, as provided in, *e.g.*, Clark *et al.*, (1994) *J. Biol. Chem.*, 269:28314-22; Mizutani *et al.*, (1997) *Life Sci.*, 61:1497-506; Rust *et al.*, (1998) *J. Mol. Endocrinol.*, 21(2):189-200; Fleury *et al.*, (1998) *J. Mol. Endocrinol.*, 21(2):131-9; Pilon *et al.*, (1997) *Endocrinology*, 138:1085-91; Lee *et al.*, (1997) *Biochem. Biophys. Res. Commun.*, 230(3):528-32; Hartung *et al.*, (1995)

Biochem. Biophys. Res. Commun., 215(2):646-53; and, *e.g.*, in GenBank Accession Nos: U66490, U53020, AB001349, L36062, L36061, S73839, AF031696, AF086814, Y17260, AB006007, AF038553, and NM_000349, or fragments thereof, or which specifically binds to polyclonal antibodies generated against any of these proteins. In preferred embodiments, a human StAR polypeptide will comprise at least about 50%, 70%, 80%, 90%, or more amino acid sequence identity to SEQ ID NO:2. Typically, the human StAR polypeptide will be expressed from the endogenous StAR gene located at genomic map position 8p11.2. A "StAR protein" or "polypeptide" can comprise naturally occurring or synthetic amino acids, *e.g.*, labeled or otherwise modified amino acids or amino acid analogs. A "StAR protein" will typically contain one or more characteristic protein motifs, any of which can be used independently of other elements normally present in a full-length StAR protein, and will have one or more characteristic activities or properties, *e.g.*, mitochondrial localization, phosphorylation patterns, or cholesterol translocating activity. A "StAR protein" can refer to any naturally occurring or synthetic StAR polypeptide as described above.

The "level of StAR mRNA" in a biological sample refers to the amount of mRNA transcribed from an StAR gene that is present in a cell or a biological sample. The mRNA generally encodes a functional StAR protein, although mutations or microdeletions may be present that alter or eliminate the function of the encoded protein. A "level of StAR mRNA" need not be quantified, but can simply be detected, *e.g.*, a subjective, visual detection by a human, with or without comparison to a level from a control sample or a level expected of a control sample.

The "level of StAR protein or polypeptide" in a biological sample refers to the amount of polypeptide translated from a StAR mRNA that is present in a cell or biological sample. The polypeptide may or may not have StAR protein activity. A "level of StAR protein" need not be quantified, but can simply be detected, *e.g.*, a subjective, visual detection by a human, with or without comparison to a level from a control sample or a level expected of a control sample.

A "full length" StAR protein or nucleic acid refers to a StAR polypeptide or polynucleotide sequence, or a variant thereof, that contains all of the elements normally contained in one or more naturally occurring, wild type StAR polynucleotide or polypeptide sequences.

As used herein, a "nucleic acid probe or oligonucleotide" is defined as a nucleic acid capable of binding to a target nucleic acid of complementary sequence through one or more types of chemical bonds, usually through complementary base pairing, usually through hydrogen bond formation. As used herein, a probe may include natural (*i.e.*, A, G, C, or T) or modified bases (7-deazaguanosine, inosine, *etc.*). In addition, the bases in a probe may be joined by a linkage other than a phosphodiester bond, so long as it does not interfere with hybridization. Thus, for example, probes may be peptide nucleic acids in which the constituent bases are joined by peptide bonds rather than phosphodiester linkages. It will be understood by one of skill in the art that probes may bind target sequences lacking complete complementarity with the probe sequence depending upon the stringency of the hybridization conditions. The probes are preferably directly labeled as with isotopes, chromophores, lumiphores, chromogens, or indirectly labeled such as with biotin to which a streptavidin complex may later bind. By assaying for the presence or absence of the probe, one can detect the presence or absence of the select sequence or subsequence.

When a quantified level of StAR falls outside of a given confidence interval for a normal level of StAR, the difference between the two levels is said to be "statistically significant." If a test value falls outside of a given confidence interval for a normal level of StAR, it is possible to calculate the probability that the test value is truly abnormal and does not simply represent a normal deviation from the average. In the present invention, a difference between a test sample and a control can be termed "statistically significant" when the probability of the test sample being a normal deviation from the average can be any of a number of values, including 0.15, 0.1, 0.05, and 0.01. Numerous sources teach how to assess statistical significance, such as Freund, J.E. (1988) *Modern elementary statistics*, Prentice-Hall.

The terms "identical" or percent "identity," in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same, when compared and aligned for maximum correspondence over a comparison window or designated region, as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection.

The phrase "substantially identical," in the context of two nucleic acids or polypeptides, refers to two or more sequences or subsequences that have at least 60%,

preferably 80%, most preferably 90-95% nucleotide or amino acid residue identity, when compared and aligned for maximum correspondence, as measured using one of the following sequence comparison algorithms or by visual inspection. Preferably, the substantial identity exists over a region of the sequences that is at least about 50 residues in length, more preferably over a region of at least about 100 residues, and most preferably the sequences are substantially identical over at least about 150 residues. In a most preferred embodiment, the sequences are substantially identical over the entire length of the coding regions.

For sequence comparison, typically one sequence acts as a reference sequence, to which test sequences are compared. When using a sequence comparison algorithm, test and reference sequences are entered into a computer, subsequence coordinates are designated, if necessary, and sequence algorithm program parameters are designated. Default program parameters can be used, or alternative parameters can be designated. The sequence comparison algorithm then calculates the percent sequence identities for the test sequences relative to the reference sequence, based on the program parameters.

A "comparison window", as used herein, includes reference to a segment of any one of the number of contiguous positions selected from the group consisting of from 20 to 600, usually about 50 to about 200, more usually about 100 to about 150 in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned. Methods of alignment of sequences for comparison are well-known in the art. Optimal alignment of sequences for comparison can be conducted, *e.g.*, by the local homology algorithm of Smith & Waterman, *Adv. Appl. Math.*, 2:482 (1981), by the homology alignment algorithm of Needleman & Wunsch, *J. Mol. Biol.*, 48:443 (1970), by the search for similarity method of Pearson & Lipman, *Proc. Natl. Acad. Sci. USA*, 85:2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by manual alignment and visual inspection (*see, e.g., Current Protocols in Molecular Biology* (Ausubel *et al.*, eds. 1995 supplement)).

One example of a useful algorithm is PILEUP. PILEUP creates a multiple sequence alignment from a group of related sequences using progressive, pairwise alignments to show relationship and percent sequence identity. It also plots a tree or

dendogram showing the clustering relationships used to create the alignment. PILEUP uses a simplification of the progressive alignment method of Feng & Doolittle, *J. Mol. Evol.*, 35:351-360 (1987). The method used is similar to the method described by Higgins & Sharp, *CABIOS* 5:151-153 (1989). The program can align up to 300 sequences, each of a maximum length of 5,000 nucleotides or amino acids. The multiple alignment procedure begins with the pairwise alignment of the two most similar sequences, producing a cluster of two aligned sequences. This cluster is then aligned to the next most related sequence or cluster of aligned sequences. Two clusters of sequences are aligned by a simple extension of the pairwise alignment of two individual sequences. The final alignment is achieved by a series of progressive, pairwise alignments. The program is run by designating specific sequences and their amino acid or nucleotide coordinates for regions of sequence comparison and by designating the program parameters. Using PILEUP, a reference sequence is compared to other test sequences to determine the percent sequence identity relationship using the following parameters: default gap weight (3.00), default gap length weight (0.10), and weighted end gaps. PILEUP can be obtained from the GCG sequence analysis software package, *e.g.*, version 7.0 (Devereaux *et al.*, *Nuc. Acids Res.*, 12:387-395 (1984).

Another example of algorithm that is suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.*, *Nuc. Acids Res.*, 25:3389-3402 (1977) and Altschul *et al.*, *J. Mol. Biol.*, 215:403-410 (1990), respectively. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). This algorithm involves first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database sequence. T is referred to as the neighborhood word score threshold (Altschul *et al.*, *supra*). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always > 0) and N (penalty score for mismatching residues; always < 0). For amino acid sequences, a scoring matrix is used to calculate the

cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) or 10, M=5, N=-4 and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength of 3, and expectation (E) of 10, and the BLOSUM62 scoring matrix (*see* Henikoff & Henikoff, *Proc. Natl. Acad. Sci. USA,* 89:10915 (1989)) alignments (B) of 50, expectation (E) of 10, M=5, N=-4, and a comparison of both strands.

The BLAST algorithm also performs a statistical analysis of the similarity between two sequences (*see, e.g.*, Karlin & Altschul, *Proc. Natl. Acad. Sci. USA,* 90:5873-5787 (1993)). One measure of similarity provided by the BLAST algorithm is the smallest sum probability (P(N)), which provides an indication of the probability by which a match between two nucleotide or amino acid sequences would occur by chance. For example, a nucleic acid is considered similar to a reference sequence if the smallest sum probability in a comparison of the test nucleic acid to the reference nucleic acid is less than about 0.2, more preferably less than about 0.01, and most preferably less than about 0.001.

An indication that two nucleic acid sequences or polypeptides are substantially identical is that the polypeptide encoded by the first nucleic acid is immunologically cross reactive with the antibodies raised against the polypeptide encoded by the second nucleic acid, as described below. Thus, a polypeptide is typically substantially identical to a second polypeptide, for example, where the two peptides differ only by conservative substitutions. Another indication that two nucleic acid sequences are substantially identical is that the two molecules or their complements hybridize to each other under stringent conditions, as described below. Yet another indication that two nucleic acid sequences are substantially identical is that the same primers can be used to amplify the sequence.

The phrase "selectively (or specifically) hybridizes to" refers to the binding, duplexing, or hybridizing of a molecule only to a particular nucleotide sequence under

stringent hybridization conditions when that sequence is present in a complex mixture (*e.g.*, total cellular or library DNA or RNA).

The phrase "stringent hybridization conditions" refers to conditions under which a probe will hybridize to its target subsequence, typically in a complex mixture of nucleic acid, but to no other sequences. Stringent conditions are sequence-dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Techniques in Biochemistry and Molecular Biology--Hybridization with Nucleic Probes*, "Overview of principles of hybridization and the strategy of nucleic acid assays" (1993). Generally, stringent conditions are selected to be about 5-10°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength pH. The T_m is the temperature (under defined ionic strength, pH, and nucleic concentration) at which 50% of the probes complementary to the target hybridize to the target sequence at equilibrium (as the target sequences are present in excess, at T_m, 50% of the probes are occupied at equilibrium). Stringent conditions will be those in which the salt concentration is less than about 1.0 M sodium ion, typically about 0.01 to 1.0 M sodium ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30°C for short probes (*e.g.*, 10 to 50 nucleotides) and at least about 60°C for long probes (*e.g.*, greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. For high stringency hybridization, a positive signal is at least two times background, preferably 10 times background hybridization. Exemplary high stringency or stringent hybridization conditions include: 50% formamide, 5x SSC and 1% SDS incubated at 42°C or 5x SSC and 1% SDS incubated at 65°C, with a wash in 0.2x SSC and 0.1% SDS at 65°C. Washes can be performed, *e.g.*, for 2, 5, 10, 15, 30, 60, or more minutes.

Nucleic acids that do not hybridize to each other under stringent conditions are still substantially identical if the polypeptides that they encode are substantially identical. This occurs, for example, when a copy of a nucleic acid is created using the maximum codon degeneracy permitted by the genetic code. In such cased, the nucleic acids typically hybridize under moderately stringent hybridization conditions. Exemplary "moderately stringent hybridization conditions" include a hybridization in a buffer of 40% formamide, 1 M NaCl, 1% SDS at 37°C, and a wash in 1X SSC at 45°C. A positive hybridization is at

least twice background. Those of ordinary skill will readily recognize that alternative hybridization and wash conditions can be utilized to provide conditions of similar stringency.

"Antibody" refers to a polypeptide comprising a framework region from an immunoglobulin gene or fragments thereof that specifically binds and recognizes an antigen. The recognized immunoglobulin genes include the kappa, lambda, alpha, gamma, delta, epsilon, and mu constant region genes, as well as the myriad immunoglobulin variable region genes. Light chains are classified as either kappa or lambda. Heavy chains are classified as gamma, mu, alpha, delta, or epsilon, which in turn define the immunoglobulin classes, IgG, IgM, IgA, IgD and IgE, respectively.

An exemplary immunoglobulin (antibody) structural unit comprises a tetramer. Each tetramer is composed of two identical pairs of polypeptide chains, each pair having one "light" (about 25 kD) and one "heavy" chain (about 50-70 kD). The N-terminus of each chain defines a variable region of about 100 to 110 or more amino acids primarily responsible for antigen recognition. The terms variable light chain (V_L) and variable heavy chain (V_H) refer to these light and heavy chains respectively.

Antibodies exist, *e.g.*, as intact immunoglobulins or as a number of well-characterized fragments produced by digestion with various peptidases. Thus, for example, pepsin digests an antibody below the disulfide linkages in the hinge region to produce $F(ab)'_2$, a dimer of Fab which itself is a light chain joined to V_H-C_H1 by a disulfide bond. The $F(ab)'_2$ may be reduced under mild conditions to break the disulfide linkage in the hinge region, thereby converting the $F(ab)'_2$ dimer into an Fab' monomer. The Fab' monomer is essentially Fab with part of the hinge region (*see, Fundamental Immunology* (Paul ed., 3d ed. 1993). While various antibody fragments are defined in terms of the digestion of an intact antibody, one of skill will appreciate that such fragments may be synthesized *de novo* either chemically or by using recombinant DNA methodology. Thus, the term antibody, as used herein, also includes antibody fragments either produced by the modification of whole antibodies, or those synthesized *de novo* using recombinant DNA methodologies (*e.g.*, single chain Fv) or those identified using phage display libraries (*see, e.g.*, McCafferty *et al.*, *Nature* 348:552-554 (1990))

For preparation of monoclonal or polyclonal antibodies, any technique known in the art can be used (*see, e.g.*, Kohler & Milstein, *Nature*, 256:495-497 (1975); Kozbor *et*

al., *Immunology Today*, 4:72 (1983); Cole *et al.*, pp. 77-96 in *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc. (1985)). Techniques for the production of single chain antibodies (U.S. Patent 4,946,778) can be adapted to produce antibodies to polypeptides of this invention. Also, transgenic mice, or other organisms such as other mammals, may be used to express humanized antibodies. Alternatively, phage display technology can be used to identify antibodies and heteromeric Fab fragments that specifically bind to selected antigens (*see*, *e.g.*, McCafferty *et al.*, *Nature*, 348:552-554 (1990); Marks *et al.*, *Biotechnology*, 10:779-783 (1992)).

A "chimeric antibody" is an antibody molecule in which (a) the constant region, or a portion thereof, is altered, replaced or exchanged so that the antigen binding site (variable region) is linked to a constant region of a different or altered class, effector function and/or species, or an entirely different molecule which confers new properties to the chimeric antibody, *e.g.*, an enzyme, toxin, hormone, growth factor, drug, etc.; or (b) the variable region, or a portion thereof, is altered, replaced or exchanged with a variable region having a different or altered antigen specificity.

The term "immunoassay" is an assay that uses an antibody to specifically bind an antigen. The immunoassay is characterized by the use of specific binding properties of a particular antibody to isolate, target, detect, and/or quantify the antigen.

The phrase "specifically (or selectively) binds" to an antibody or "specifically (or selectively) immunoreactive with," when referring to a protein or peptide, refers to a binding reaction that is determinative of the presence of the protein in a heterogeneous population of proteins and other biologics. Thus, under designated immunoassay conditions, the specified antibodies bind to a particular protein at least two times the background and do not substantially bind in a significant amount to other proteins present in the sample. Specific binding to an antibody under such conditions may require an antibody that is selected for its specificity for a particular protein. For example, polyclonal antibodies raised to a StAR polypeptide can be selected to obtain only those polyclonal antibodies that are specifically immunoreactive with StAR and not with other proteins, except for polymorphic variants, orthologs, and alleles of StAR. A variety of immunoassay formats may be used to select antibodies specifically immunoreactive with a particular protein. For example, solid-phase ELISA immunoassays are routinely used to select antibodies specifically immunoreactive with a protein (*see, e.g.*, Harlow & Lane, *Antibodies, A Laboratory Manual*

(1988) for a description of immunoassay formats and conditions that can be used to determine specific immunoreactivity). Typically a specific or selective reaction will be at least twice background signal or noise and more typically more than 10 to 100 times background.

The phrase "selectively associates with" refers to the ability of a nucleic acid to "selectively hybridize" with another as defined above, or the ability of an antibody to "selectively (or specifically) bind to a protein, as defined above.

III. Detecting StAR Polynucleotides and Polypeptides

The present invention is based on the discovery that a diagnosis and/or prognosis of cancer can be determined by detecting elevated levels of StAR polypeptide or polynucleotides in a biological sample from a mammal. As described below, any of a number of methods to detect the presence or level of StAR can be used. A StAR polynucleotide level can be detected by detecting any StAR DNA or RNA, including StAR genomic DNA, mRNA, and cDNA. A StAR polypeptide can be detected by detecting an StAR polypeptide itself, or by detecting StAR protein activity, *e.g.*, cholesterol translocation, steroid production, *etc.* Detection can involve quantification of the level of StAR (*e.g.*, gDNA, cDNA, mRNA, or protein level, or protein activity), or, alternatively, can be a qualitative assessment of the level, or of the presence or absence, of StAR, in particular in comparison with a control level. Any of a number of methods to detect any of the above can be used, as described *infra*. Such methods include, for example, hybridization, amplification, and other assays.

In certain embodiments, a level of StAR in a biological sample will be compared with a control sample taken from a cancer-free animal, or, preferably, with a value expected for a sample taken from a cancer-free animal. In a particularly preferred embodiment, an assay will be performed under conditions where only a higher than normal amount of StAR polynucleotide or polypeptide will be detectable in the assay. As a result, an elevated level of StAR can be detected in a sample using a simple assay giving a simple, positive or negative result, with no need for quantification of StAR levels or direct comparison with a control level.

In certain embodiments, the level of StAR polynucleotide, polypeptide, or protein activity will be quantified. In such embodiments, the difference between an elevated

level of StAR and a normal, control level will preferably be statistically significant. In preferred embodiments, an elevated level of StAR polynucleotide, polypeptide, and/or protein activity will be at least about 2, 5, 10, or more fold greater than a control level, or a level expected of a control sample.

In certain embodiments, the ability to detect an increased level, or diagnostic presence, in a cell is used as a marker for cancer cells, *i.e.*, to monitor the number or localization of cancer cells in a patient, as detected *in vivo* or *in vitro*.

Typically, the StAR polynucleotides or polypeptides detected herein will be at least about 70% identical, and preferably 80% or more identical, over a region of at least about 50, 100, 200, or more nucleotides, or 20, 50, 100, or more amino acids, to SEQ ID NO:1 or 2, or to one or more StAR sequences available, *e.g.*, from GenBank (*see, e.g.*, GenBank Accession Nos: U66490, U53020, AB001349, L36062, L36061, S73839, AF031696, AF086814, Y17260, AB006007, AF038553, and NM_000349). Such polynucleotides or polypeptides can represent functional or nonfunctional forms of StAR, or any variant, derivative, or fragment thereof.

Typically, the level and/or presence of StAR polynucleotides or polypeptides will be detected in a biological sample. A "biological sample" refers to a cell or population of cells or a quantity of tissue or fluid from an animal. Most often, the sample has been removed from an animal, but the term "biological sample" can also refer to cells or tissue analyzed *in vivo*, *i.e.*, without removal from the animal. Typically, a "biological sample" will contain cells from the animal, but the term can also refer to noncellular biological material, such as noncellular fractions of blood, saliva, or urine, that can be used to measure StAR levels. Numerous types of biological samples can be used in the present invention, including, but not limited to, a tissue biopsy, blood sample, a buccal scrape, a saliva sample, or a nipple discharge.

As used herein, a "tissue biopsy" refers to an amount of tissue removed from an animal for diagnostic analysis. In a patient with cancer, tissue may be removed from a tumor, allowing the analysis of cells within the tumor. "Tissue biopsy" can refer to any type of biopsy, such as needle biopsy, fine needle biopsy, surgical biopsy, *etc.* A "buccal scrape" refers to a sample of cells removed from the inner lining of the mouth. A "nipple discharge" refers to fluid originating from a nipple, which may contain cancerous cells or may contain elevated levels of StAR polypeptide indicating the presence of cancerous cells in the breast.

A. DETECTION OF COPY NUMBER

In one embodiment, the presence of cancer is evaluated by determining the copy number of StAR genes, *i.e.*, the number of DNA sequences in a cell encoding a StAR protein. Generally, for a given autosomal gene, an animal has two copies of each gene. The copy number can be increased, however, by gene amplification or duplication, *e.g.*, in cancer cells, or reduced by deletion. Methods of evaluating the copy number of a particular gene are well known to those of skill in the art, and include, *inter alia*, hybridization and amplification based assays.

1. Hybridization-based Assays

Any of a number of hybridization based assays can be used to detect the copy number of StAR genes in the cells of a biological sample. One such method is by Southern blot. In a Southern blot, genomic DNA is typically fragmented, separated electrophoretically, transferred to a membrane, and subsequently hybridized to a StAR specific probe. Comparison of the intensity of the hybridization signal from the probe for the target region with a signal from a control probe for a region of normal genomic DNA (*e.g.*, a nonamplified portion of the same or related cell, tissue, organ, *etc.*) provides an estimate of the relative StAR copy number. Southern blot methodology is well known in the art and is described, *e.g.*, in Ausubel *et al.*, or Sambrook *et al.*, *supra*.

An alternative means for determining the copy number of StAR genes in a sample is by *in situ* hybridization, *e.g.*, fluorescence *in situ* hybridization, or FISH. *In situ* hybridization assays are well known (*e.g.*, Angerer (1987) *Meth. Enzymol.*, 152:649). Generally, *in situ* hybridization comprises the following major steps: (1) fixation of tissue or biological structure to be analyzed; (2) prehybridization treatment of the biological structure to increase accessibility of target DNA, and to reduce nonspecific binding; (3) hybridization of the mixture of nucleic acids to the nucleic acid in the biological structure or tissue; (4) post-hybridization washes to remove nucleic acid fragments not bound in the hybridization and (5) detection of the hybridized nucleic acid fragments.

The probes used in such applications are typically labeled, *e.g.*, with radioisotopes or fluorescent reporters. Preferred probes are sufficiently long, *e.g.*, from about 50, 100, or 200 nucleotides to about 1000 or more nucleotides, so as to specifically hybridize with the target nucleic acid(s) under stringent conditions.

In numerous embodiments, "comparative probe" methods, such as comparative genomic hybridization (CGH), are used to detect StAR gene amplification. In comparative genomic hybridization methods, a "test" collection of nucleic acids is labeled with a first label, while a second collection (*e.g.*, from a healthy cell or tissue) is labeled with a second label. The ratio of hybridization of the nucleic acids is determined by the ratio of the first and second labels binding to each fiber in an array. Differences in the ratio of the signals from the two labels, *e.g.*, due to gene amplification in the test collection, is detected and the ratio provides a measure of the StAR gene copy number.

Hybridization protocols suitable for use with the methods of the invention are described, *e.g.*, in Albertson (1984) *EMBO J.*, 3:1227-1234; Pinkel (1988) *Proc. Natl. Acad. Sci. USA*, 85:9138-9142; EPO Pub. No. 430,402; *Methods in Molecular Biology, Vol. 33: In Situ Hybridization Protocols*, Choo, ed., Humana Press, Totowa, NJ (1994), *etc.*

2. Amplification-based Assays

In another embodiment, amplification-based assays are used to measure the copy number of StAR genes. In such assays, the StAR nucleic acid sequences act as a template in an amplification reaction (*e.g.*, Polymerase Chain Reaction, or PCR). In a quantitative amplification, the amount of amplification product will be proportional to the amount of template in the original sample. Comparison to appropriate controls provides a measure of the copy number of the StAR gene. Methods of quantitative amplification are well known to those of skill in the art. Detailed protocols for quantitative PCR are provided, *e.g.*, in Innis *et al.* (1990) *PCR Protocols, A Guide to Methods and Applications*, Academic Press, Inc. N.Y.). The known nucleic acid sequence for StAR (*see, e.g.*, SEQ ID NO:1) is sufficient to enable one of skill to routinely select primers to amplify any portion of the gene.

In preferred embodiments, a TaqMan based assay is used to quantify StAR polynucleotides. TaqMan based assays use a fluorogenic oligonucleotide probe that contains a 5' fluorescent dye and a 3' quenching agent. The probe hybridizes to a PCR product, but cannot itself be extended due to a blocking agent at the 3' end. When the PCR product is amplified in subsequent cycles, the 5' nuclease activity of the polymerase, *e.g.*, AmpliTaq, results in the cleavage of the TaqMan probe. This cleavage separates the 5' fluorescent dye and the 3' quenching agent, thereby resulting in an increase in fluorescence as a function of amplification (*see*, for example, literature provided by Perkin-Elmer, *e.g.*, www2.perkin-elmer.com).

Other suitable amplification methods include, but are not limited to, ligase chain reaction (LCR) (*see,* Wu and Wallace (1989) *Genomics,* 4:560, Landegren *et al.* (1988) *Science,* 241:1077, and Barringer *et al.* (1990) *Gene,* 89:117), transcription amplification (Kwoh *et al.* (1989) *Proc. Natl. Acad. Sci. USA,* 86:1173), self-sustained sequence replication (Guatelli *et al.* (1990) *Proc. Nat. Acad. Sci. USA,* 87:1874), dot PCR, and linker adapter PCR, *etc.*

B. DETECTION OF GENE EXPRESSION

In preferred embodiments, StAR levels are characterized by detecting StAR gene expression by virtue of levels of StAR mRNA, StAR polypeptide and/or StAR protein activity (*e.g.*, cholesterol translocation activity) in a biological sample.

1. Detection of StAR mRNA

a) Direct hybridization-based assays

Methods of detecting and/or quantifying the level of StAR gene transcript (StAR mRNA or cDNA made therefrom) using nucleic acid hybridization techniques are known to those of skill in the art (*see*, Sambrook *et al.*, (1989) *Molecular Cloning: A Laboratory Manual*, 2d Ed., vols 1-3, Cold Spring Harbor Press, New York).

For example, one method for evaluating the presence, absence, or quantity of StAR cDNA involves a Southern Blot as described above. Briefly, StAR mRNA is isolated using standard methods and reverse transcribed to produce cDNA. The cDNA is then optionally digested, run on a gel, and transferred to a membrane. Hybridization is then carried out using nucleic acid probes specific for StAR cDNA and detected using standard techniques (*see, e.g.*, Sambrook *et al.*, *supra*).

Similarly, a Northern blot may be used to detect an mRNA directly. In brief, in a typical embodiment, mRNA is isolated from a given biological sample, electrophoresed to separate the mRNA species, and transferred from the gel to a nitrocellulose membrane. As with the Southern blots, labeled StAR probes are then hybridized to the membrane to identify and/or quantify the mRNA.

b) Amplification-based assays

In another preferred embodiment, a StAR transcript (*e.g.*, StAR mRNA) is detected using amplification-based methods (*e.g.*, RT-PCR). RT-PCR methods are well known to those of skill (*see, e.g.*, Ausubel *et al., supra*). Preferably, quantitative RT-PCR is used, thereby allowing the comparison of the level of mRNA in a sample with a control sample or value.

C. DETECTION OF EXPRESSED PROTEIN

StAR levels can also be detected and/or quantified by detecting or quantifying StAR polypeptide. StAR polypeptide can be detected and quantified by any of a number of means well known to those of skill in the art. These include analytic biochemical methods such as electrophoresis, capillary electrophoresis, high performance liquid chromatography (HPLC), thin layer chromatography (TLC), hyperdiffusion chromatography, and the like, or various immunological methods such as fluid or gel precipitin reactions, immunodiffusion (single or double), immunoelectrophoresis, radioimmunoassay (RIA), enzyme-linked immunosorbent assays (ELISAs), immunofluorescent assays, western blotting, and the like.

In a preferred embodiment, an StAR polypeptide is detected using an immunoassay such as an ELISA assay (*see, e.g.*, Crowther, John R. *ELISA Theory and Practice.* Humana Press. New Jersey, 1995). As used herein, an "immunoassay" is an assay that utilizes an antibody to specifically bind to the analyte (*i.e.*, the StAR polypeptide). The immunoassay is thus characterized by detection of specific binding of a StAR polypeptide to an anti-StAR antibody.

In an immunoassay, StAR polypeptide can be detected and/or quantified using any of a number of well recognized immunological binding assays (*see, e.g.*, U.S. Patent Nos. 4,366,241; 4,376,110; 4,517,288; and 4,837,168). For a review of the general immunoassays, see also Asai (1993) *Methods in Cell Biology,* Volume *37: Antibodies in Cell Biology,* Academic Press, Inc. New York; Stites & Terr (1991) *Basic and Clinical Immunology 7th Edition, Enzyme Immunoassay* (Maggio, ed., 1980); and Harlow & Lane, *supra.*

Immunoassays typically rely on direct or indirect labeling methods to detect antibody-analyte binding. For example, an anti-StAR antibody can be directly labeled, thereby allowing detection. Alternatively, the anti-StAR antibody may itself be unlabeled,

but may, in turn, be bound by a labeled third antibody specific to antibodies of the species from which the second antibody is derived. The second or third antibodies can also be modified with a detectable moiety, *e.g.*, as biotin, to which a third labeled molecule can specifically bind, such as enzyme-labeled streptavidin. Also, other antibody-binding molecules can be used, *e.g.*, labeled protein A or G (*see, generally* Kronval, *et al.* (1973) *J. Immunol.*, 111:1401-1406, and Akerstrom (1985) *J. Immunol.*, 135:2589-2542).

Throughout the assays, incubation and/or washing steps may be required after each combination of reagents. Incubation steps can vary from about 5 seconds to several hours, preferably from about 5 minutes to about 24 hours. However, the incubation time will depend upon the assay format, antigen, volume of solution, concentrations, and the like. Usually, the assays will be carried out at ambient temperature, although they can be conducted over a range of temperatures, such as 10°C to 40°C.

Immunoassays for detecting a StAR polypeptide can be competitive or noncompetitive. Noncompetitive immunoassays are assays in which the amount of captured analyte is directly measured. In a preferred embodiment, "sandwich" assays will be used, for example, wherein anti-StAR antibodies are bound directly to a solid substrate where they are immobilized. These immobilized antibodies then capture the StAR protein present in a test sample. The StAR thus immobilized is then bound by a labeling agent, such as a second anti-StAR antibody bearing a label.

In competitive assays, the amount of StAR protein present in a sample is measured indirectly, *e.g.*, by measuring the amount of added (exogenous) StAR displaced (or competed away) from an anti-StAR antibody by StAR protein present in a sample. For example, a known amount of labeled StAR polypeptide is added to a sample and the sample is then contacted with an anti-StAR antibody. The amount of labeled StAR polypeptide bound to the antibody is inversely proportional to the concentration of StAR polypeptide present in the sample.

Any of a number of labels can be used in any of the immunoassays of this invention, including fluorescent labels, radioisotope labels, or enzyme-based labels, wherein a detectable product of enzyme activity is detected (*e.g.*, peroxidase, alkaline phosphatase, β-galactosidase, *etc.*).

One of skill in the art will appreciate that it is often desirable to minimize nonspecific binding in immunoassays. Particularly, where the assay involves an antigen or

antibody immobilized on a solid substrate it is desirable to minimize the amount of nonspecific binding to the substrate. Means of reducing such nonspecific binding are well known to those of skill in the art. Typically, this technique involves coating the substrate with a proteinaceous composition. In particular, protein compositions such as bovine serum albumin (BSA), nonfat powdered milk, and gelatin are widely used.

Methods of producing polyclonal and monoclonal antibodies that react specifically with StAR proteins are known to those of skill in the art (*see, e.g.*, Coligan, *Current Protocols in Immunology* (1991); Harlow & Lane, *supra*; Goding, *Monoclonal Antibodies: Principles and Practice* (2d ed. 1986); and Kohler & Milstein, *Nature,* 256:495-497 (1975). Such techniques include antibody preparation by selection of antibodies from libraries of recombinant antibodies in phage or similar vectors, as well as preparation of polyclonal and monoclonal antibodies by immunizing rabbits or mice (*see, e.g.*, Huse *et al.*, *Science,* 246:1275-1281 (1989); Ward *et al.*, *Nature,* 341:544-546 (1989)).

A number of StAR peptides or a full length protein may be used to produce antibodies specifically reactive with StAR protein. For example, recombinant StAR protein can be expressed in eukaryotic or prokaryotic cells and purified using standard methods. Recombinant protein is the preferred immunogen for the production of monoclonal or polyclonal antibodies. Alternatively, a synthetic peptide derived from any StAR amino acid sequence can be conjugated to a carrier protein and used as an immunogen. Naturally occurring protein may also be used either in pure or impure form. The product is then injected into an animal capable of producing antibodies. Either monoclonal or polyclonal antibodies may be generated, for subsequent use in immunoassays to measure the protein.

Methods of production of polyclonal antibodies are known to those of skill in the art. An inbred strain of mice (*e.g.*, BALB/C mice) or rabbits is immunized with the protein using a standard adjuvant, such as Freund's adjuvant, and a standard immunization protocol. The animal's immune response to the immunogen preparation is monitored by taking test bleeds and determining the titer of reactivity to the StAR protein. When appropriately high titers of antibody to the immunogen are obtained, blood is collected from the animal and antisera are prepared. Further fractionation of the antisera to enrich for antibodies reactive to the protein can be done if desired (*see,* Harlow & Lane, *supra*).

Monoclonal antibodies may be obtained by various techniques familiar to those skilled in the art. Briefly, spleen cells from an animal immunized with a desired

antigen are immortalized, commonly by fusion with a myeloma cell (*see*, Kohler & Milstein, *Eur. J. Immunol.*, 6:511-519 (1976)). Alternative methods of immortalization include transformation with Epstein Barr Virus, oncogenes, or retroviruses, or other methods well known in the art. Colonies arising from single immortalized cells are screened for production of antibodies of the desired specificity and affinity for the antigen, and yield of the monoclonal antibodies produced by such cells may be enhanced by various techniques, including injection into the peritoneal cavity of a vertebrate host. Alternatively, one may isolate DNA sequences which encode a monoclonal antibody or a binding fragment thereof by screening a DNA library from human B cells according to the general protocol outlined by Huse *et al.*, *Science*, 246:1275-1281 (1989).

Monoclonal antibodies and polyclonal sera are collected and titered against the immunogen protein in an immunoassay, for example, a solid phase immunoassay with the immunogen immobilized on a solid support. Typically, polyclonal antisera with a titer of 10^4 or greater are selected and tested for their cross reactivity against non-StAR proteins or even other related proteins from other organisms, using a competitive binding immunoassay. Specific polyclonal antisera and monoclonal antibodies will usually bind with a K_d of at least about 0.1 mM, more usually at least about 1 μM, preferably at least about 0.1 μM or better, and most preferably, 0.01 μM or better.

D. DETECTION OF STAR PROTEIN ACTIVITY OR BEHAVIOR

In another embodiment, StAR polypeptide levels are determined by virtue of the StAR protein activity or behavior in a biological sample. Such protein activity or behavior can be easily measured using standard techniques. For example, the StAR protein regulates the translocation of cholesterol from the outer to the inner mitochondrial membrane, the rate-limiting step in steroid biosynthesis. Thus, StAR protein activity can be indirectly detected in a biological sample by detecting the amount of cholesterol translocation, or steroid production, that is present in the sample. Such assays are well known to those of skill in the art.

IV. Diagnosing Cancer

The present invention provides numerous methods for diagnosing any of a number of types of cancer, *e.g.*, determining whether or not a mammal has cancer, whether

or not a biological sample contains cancerous cells, estimating the likelihood of a mammal developing cancer, and monitoring the efficacy of anti-cancer treatment in a mammal with cancer. Such methods are based on the discovery that cancer cells have an elevated level of StAR polynucleotide (*i.e.*, gene copy number and/or mRNA) and polypeptide level. Accordingly, by determining whether or not a cell contains elevated levels of StAR polynucleotide or polypeptide, it is possible to determine whether or not the cell is cancerous. Further, the presence of cancerous cells can be determined indirectly, *i.e.*, in certain embodiments a biological sample that does not itself contain cancerous cells, but which has been taken from an animal with cancerous cells elsewhere in its body, may contain elevated levels of StAR reflecting the presence of the cancerous cells.

A. DETECTING A CANCER

In numerous embodiments of the present invention, the level and/or presence or StAR polynucleotide or polypeptide will be detected in a biological sample, thereby detecting the presence or absence of cancerous cells in the biological sample, or, in certain embodiments, in the mammal from which the biological sample was removed. In preferred embodiments, the biological sample will comprise a tissue sample from a tissue suspected of containing cancerous cells. For example, in a woman suspected of having breast cancer, breast tissue is removed. Often, such methods will be used in conjunction with additional diagnostic methods, *e.g.*, detection of other cancer markers, mammography, *etc.* In other embodiments, a tissue sample known to contain cancerous cells, *e.g.*, from a tumor, will be detected for StAR levels to determine information about the cancer, *e.g.*, the efficacy of certain treatments, the survival expectancy of the animal, *etc.*

The amount of StAR polynucleotide or polypeptide that will indicate the presence of a cancer will depend on numerous factors, including the type of cancer, the age, sex, medical history, *etc.*, of the patient, the cell type, the assay format, *etc.* In preferred embodiments, a level of StAR in a biological sample will not be quantified or directly compared with a control sample, but will rather be detected relative to a "diagnostic presence" of StAR, wherein a "diagnostic presence" refers to an amount of StAR polynucleotide or polypeptide that indicates the presence of cancer, or indicates a likelihood of cancer, in the mammal from which the sample was taken. Preferably, a "diagnostic presence" will be detectable in a simple assay giving a positive or negative result, where a

positive "detection" of a "diagnostic presence" of StAR polynucleotide or polypeptide indicates the presence of cancer in the mammal.

The StAR level need not be quantified for a "diagnostic presence" to be detected, merely any method of determining whether StAR is present at levels higher than in a normal, cancer-free cell, sample, or mamml. In addition, a "diagnostic presence" does not refer to any absolute quantity of StAR, but rather on an amount that, depending on the biological sample, cell type, assay conditions, medical condition of the mammal, *etc.*, is sufficient to distinguish the level in a cancerous, or pre-cancerous sample, from a normal, cancer-free sample.

Such methods can be practiced regardless of whether any StAR polynucleotide or polypeptide is normally present, or "expected" to be present, in a particular control sample. For example, StAR may not be expressed in certain cell types, resulting in a complete absence of StAR in a control biological sample consisting of such cell types. For such biological sample, a "diagnostic presence" refers to any detectable amount of StAR, using any assay. In other tissues, however, there may be a detectable level of StAR present in normal, cancer-free cells, and a "diagnostic presence" represents a level that is higher than the normal level, preferably representing a "statistically significant" increase over the normal level. Often, a "diagnostic presence" of StAR polynucleotide, polypeptide, and/or protein activity in a biological sample will be at least about 1.5, 2, 5, 10, or more fold greater than a level expected in a sample taken from a normal, cancer-free mammal.

Further, the present methods can be used to assess the efficacy of a course of treatment. For example, in a mammal with cancer from which a biological sample has been found to contain an elevated amount of StAR polynucleotide or polypeptide, the efficacy of an anti-cancer treatment can be assessed by monitoring, over time, StAR levels. For example, a reduction in StAR polynucleotide or polypeptide levels in a biological sample taken from a mammal following a treatment, compared to a level in a sample taken from the mammal before, or earlier in, the treatment, indicates efficacious treatment.

B. DETERMINING A PROGNOSIS

The level of StAR can be used to determine the prognosis of a mammal with cancer. For example, if cancer is detected using a technique other than by detecting StAR,

e.g., tissue biopsy, then the presence or absence of StAR can be used to determine the prognosis for the mammal, *i.e.*, an elevated level of StAR will indicate a reduced survival expectancy in the mammal compared to in a mammal with cancer but with a normal level of StAR. As used herein, "survival expectancy" refers to a prediction regarding the severity, duration, or progress of a disease, condition, or any symptom thereof. In a preferred embodiment, an increased level, a diagnostic presence, or a quantified level, of StAR is statistically correlated with the observed progress of a disease, condition, or symptom in a large number of mammals, thereby providing a database wherefrom a statistically-based prognosis can be made in view of any detected level or presence of StAR. For example, in a particular type of mammal, *i.e.*, a human of a particular age, gender, medical condition, medical history, *etc.*, a detection of a level of StAR that is, *e.g.*, 2 fold higher than a control level may indicate, *e.g.*, a 10% reduced survival expectancy in the human compared to in a similar human with a normal level of StAR, based on a previous study of the level of StAR in a large number of similar patients whose disease progression was observed and recorded.

C. DETERMINING A PREFERRED COURSE OF TREATMENT

The present methods can be used to determine the optimal course of treatment in a mammal with cancer. For example, the presence of an elevated level of StAR can indicate a reduced survival expectancy of a mammal with cancer, thereby indicating a more aggressive treatment for the mammal. In addition, a correlation can be readily established between levels of StAR, or the presence or absence of a diagnostic presence of StAR, and the relative efficacy of one or another anti-cancer agent. Such analyses can be performed, *e.g.*, retrospectively, *i.e.*, by detecting StAR levels in samples taken previously from mammals that have subsequently undergone one or more types of anti-cancer therapy, and correlating the StAR levels with the known efficacy of the treatment.

In numerous embodiments, levels of StAR polynucleotides or polypeptides in tumor cells of a mammal, *e.g.*, as detected by immunoassay using anti-StAR antibodies, are used to guide the selection of an anti-cancer treatment based on the effects of the treatment on steroid hormone levels or activity. For example, StAR protein levels, as detected by anti-StAR antibodies, in breast tumors can be used to predict the response of a patient to tamoxifen or any other treatment that affects estrogen metabolism or activity, such as aromatase inhibitors or estrogen antagonists/agonists. Similarly, StAR protein levels, as

detected by anti-StAR antibodies, in prostate tumors can be used to predict the response of a patient to a treatment that affects testosterone metabolism or activity. In preferred embodiments, a detection of an elevated or diagnostic level of StAR indicates the beneficial use of a treatment that inhibits the activity of a steroid hormone. For example, in breast tumors, an increased or diagnostic level of StAR indicates that the use of an agent that inhibits estrogen activity, such as tamoxifen or other estrogen antagonists, an agent that inhibits estrogen production, such as an aromatase inhibitor (*e.g.*, letrozole or vorozole), or any other treatment that affects estrogen activity or levels, such as the surgical removal of the ovaries, would provide an efficacious treatment for the tumor. Similarly, in prostate tumors, an increased or diagnostic level of StAR indicates that the use of an antiandrogen such as flutamide, bicalutamide, or nilutamide, an agent that reduces testosterone levels such as a luteinizing hormone-releasing hormone (LHRH) analog, *e.g.*, *leuprolide* or *goserelin*, other hormonal drugs such as megestrol acetate, medroxyprogesterone, or ketoconazole, or any other treatment that affects testosterone activity or levels, such as orchiectomy, would provide an effective treatment for the tumor.

V. Treating cancer

The present invention provides numerous methods for treating a mammal with cancer. In addition to allowing the determination of an optimal treatment for a mammal with cancer, as described *supra*, methods are provided for treating a cancer by inhibiting the growth, proliferation, or steroid hormone production of cells within the mammal, *e.g.*, cancer cells. Typically, the methods are directed at reducing the level of StAR polypeptides, polynucleotides, or protein activity in a cancerous cell. It will be appreciated that more than one of the methods described *infra* can be performed on a given animal, and may also be administered in conjunction with one or more traditional, well known anti-cancer therapies, *e.g.*, chemotherapy, radiation therapy, surgery, hormone therapy, immunotherapy, *etc.*

According to the present invention, a "method of treating cancer" refers to a procedure or course of action that is designed to reduce or eliminate the number of cancer cells in an animal, or to alleviate the symptoms of a cancer. "A method of treating cancer" does not necessarily mean that the cancer cells will, in fact, be eliminated, that the number of cells will, in fact, be reduced, or that the symptoms of a cancer will, in fact, be alleviated. Often, a method of treating cancer will be performed even with a low likelihood of success,

but which, given the medical history and estimated survival expectancy of an animal, is nevertheless deemed an overall beneficial course of action.

In certain embodiments, the present invention provides methods for treating cancer by detecting the level and/or a diagnostic presence of StAR polynucleotide or polypeptide in a biological sample, and, when a diagnostic presence or increased level is detected, applying one or more anti-cancer therapies, including, but not limited to, chemotherapy, radiation therapy, surgery, immunotherapy, hormone therapy, and gene therapy.

One commonly applied anti-cancer therapy is chemotherapy, *i.e.*, the administration of chemical compounds to a mammal with cancer that is aimed at killing or reducing the number of cancer cells within the mammal. Generally, chemotherapeutic agents arrest the growth of or kill cells that are dividing or growing, such as cancer cells. Examples of chemotherapeutic agents include, but are not limited to, genistein, taxol, busulfan, cisplatin, cyclophosphamide (cytoxan), dacarbazine, ifosfamide, mechlorethamine (mustargen), melphalan, carmustine, lomustine, 5-fluorouracil, methotrexate, gemcitabine, cytarabine (Ara-C), fludarabine, bleomycin, dactinomycin, daunorubicin, doxorubicin (Adriamycin), idarubicin, paclitaxel, docetaxel, etoposide, vinblastine, vincristine, vinorelbine, L-asparaginase, amsacrine, tretinoin, prednisone and dexamethasone.

Another commonly applied anti-cancer therapy is radiation therapy, wherein radioactivity is administered to a mammal with cancer. Radiation kills or inhibits the growth of dividing cells, such as cancer cells. The administration of radiation may be from an external source (*e.g.*, a gamma source, a proton source, a molecular beam source, *etc.*), or may be through an implantable radioactive material, or a radioactive molecule such as an antibody.

In numerous embodiments, a tissue found to be cancerous using the present methods will be removed using surgery, *i.e.*, the direct removal or ablation of cells, *e.g.*, cancer cells, from a mammal. Most often, the cancer cells will be in the form of a tumor (*e.g.*, a mammary tumor), which is removed from the mammal. The surgical methods may involve removal of healthy as well as cancerous tissue.

Hormone therapy can also be used to treat cancers, *e.g.*, breast cancer. For example, compounds can be administered to a patient that counteract or inhibit hormones, such as estrogen or androgen, that have a mitogenic effect on cells and which often act to

increase the cancerous properties of cancer cells *in vivo*. Hormone therapy can also include methods of reducing or eliminating the production of hormones in an animal, *e.g.*, the surgical removal of ovaries in an animal to prevent estrogen production.

In certain embodiments, immunotherapy will be used to treat a cancer following a diagnosis based on detection of StAR, *i.e.*, methods of enhancing the ability of an animal's immune system to destroy cancer cells within the animal. Numerous such methods are well known to those of skill in the art. This can involve the treatment with polyclonal or monoclonal antibodies (*e.g.*, Herceptin) that bind to particular molecules located on, produced by, or indicative of, tumor cells. Immunotherapeutic methods are well know to those of skill in the art (*see, e.g.,* Pastan *et al.*(1992) *Ann. Rev. Biochem.*, 61:331-354, Brinkman and Pastan (1994) *Biochimica Biphysica Acta,* 1198:27-45, *etc.*).

In other embodiments, gene therapy will be used to treat a cancer diagnosed based on a detection of StAR. In such embodiments, a nucleic acid is introduced into cells, *e.g.*, cancer cells, to provide treatment for the cancer. For example, tumor suppressor genes that are often missing or mutated in a cancer cell, *e.g.*, p53, RB, p21, p16, and others, can be replaced or overexpressed by introducing a nucleic acid encoding a functional gene into the cells. In addition, genes whose overexpression or increased activity contributes to cancer, *e.g.*, *ras*, telomerase, *etc.*, can be inhibited by any of a number of methods, including, but not limited to, antisense, ribozymes, and polynucleotides encoding dominant negative forms or other inhibiting polypeptides. Such nucleic acids can be delivered using any of a variety of methods, *e.g.*, liposomal formulations, viral vectors, naked DNA injection, *etc.*, and can be performed *in vivo* or *ex vivo*.

A. REDUCING STAR ACTIVITY LEVELS IN CELLS

In preferred embodiments, this invention provides methods of treating a cancer by reducing StAR levels in a cell. Typically, such methods are used to reduce an elevated level of StAR, *e.g.*, an elevated level in a cancerous cell, and can be performed in any of a number of ways, *e.g.*, lowering the copy number of StAR genes or decreasing the level of StAR mRNA, protein, or protein activity in a cell. Preferably, the level of StAR activity is lowered to a level typical of a normal, cancer-free cell, but the level may be reduced to any level that is sufficient to decrease the proliferation or steroid production of the cell, including to levels above or below those typical of normal cells. Preferably, such

methods involve the use of inhibitors of StAR, where an "inhibitor of StAR" is a molecule that acts to reduce StAR polynucleotide levels, StAR polypeptide levels and/or StAR protein activity. Such inhibitor s include, but are not limited to, antisense polynucleotides, ribozymes, antibodies, dominant negative StAR forms, and small molecule inhibitors of StAR.

In preferred embodiments, StAR levels will be reduced so as to reduce the growth or proliferation of a cancer cell with elevated StAR levels. The proliferation of a cell refers to the rate at which the cell or population of cells divides, or to the extent to which the cell or population of cells divides or increases in number. Proliferation can reflect any of a number of factors, including the rate of cell growth and division and the rate of cell death. In numerous embodiments, StAR levels will be reduced so as to reduce the amount of steroid hormone production by the cell. StAR regulates the rate limiting translocation of cholesterol across the mitochondrial membranes, thereby controlling the rate of steroidogenesis in cells. Without being bound by the presently offered theory, it is suggested that the amplification and/or overexpression of the StAR gene in cancer cells, *e.g.*, breast cancer cells, causes an increased level of steroid hormone, *e.g.*, estrogen, production by the cells. This increased steroid hormone, in turn, appears to cause "autocrine" stimulation of cancer cell, *e.g.*, breast cancer cell, growth and proliferation. Thus, by decreasing the StAR activity in a cell, *e.g.*, a breast cancer cell, it is possible to reduce the amount of steroid hormone, *e.g.*, estrogen, produced by the cell, thereby inhibiting the growth of the cell itself and of neighboring cells.

The ability of any of the present compounds to affect StAR activity can be determined based on any of a number of factors, including, but not limited to, a level of StAR polynucleotide, *e.g.*, mRNA or gDNA, the level of StAR polypeptide, the degree of binding of a compound to a StAR polynucleotide or polypeptide, StAR intracellular localization, or any functional properties of StAR protein, such as the ability of StAR activity to effect cholesterol translocation into the mitochondria and the resulting steroid hormone synthesis.

1. Inhibitors of StAR polynucleotides

a) Antisense Polynucleotides

In certain embodiments, StAR activity is downregulated, or entirely inhibited, by the use of antisense polynucleotide, *i.e.*, a nucleic acid complementary to, and which can

preferably hybridize specifically to, a coding mRNA nucleic acid sequence, *e.g*, StAR mRNA, or a subsequence thereof. Binding of the antisense polynucleotide to the StAR mRNA reduces the translation and/or stability of the StAR mRNA.

In the context of this invention, antisense polynucleotides can comprise naturally-occurring nucleotides, or synthetic species formed from naturally-occurring subunits or their close homologs. Antisense polynucleotides may also have altered sugar moieties or inter-sugar linkages. Exemplary among these are the phosphorothioate and other sulfur containing species which are known for use in the art. All such analogs are comprehended by this invention so long as they function effectively to hybridize with StAR mRNA.

Such antisense polynucleotides can readily be synthesized using recombinant means, or can be synthesized *in vitro*. Equipment for such synthesis is sold by several vendors, including Applied Biosystems. The preparation of other oligonucleotides such as phosphorothioates and alkylated derivatives is also well known to those of skill in the art.

b) Ribozymes

In addition to antisense polynucleotides, ribozymes can be used to target and inhibit transcription of StAR. A ribozyme is an RNA molecule that catalytically cleaves other RNA molecules. Different kinds of ribozymes have been described, including group I ribozymes, hammerhead ribozymes, hairpin ribozymes, RNAse P, and axhead ribozymes (*see, e.g.,* Castanotto *et al.* (1994) *Adv. in Pharmacology,* 25:289-317 for a general review of the properties of different ribozymes).

The general features of hairpin ribozymes are described, *e.g.*, in Hampel *et al.* (1990) *Nucl. Acids Res.*, 18:299-304; Hampel *et al.* (1990) European Patent Publication No. 0 360 257; U.S. Patent No. 5,254,678. Methods of preparing are well known to those of skill in the art (*see, e.g.,* Wong-Staal *et al.*, WO 94/26877; Ojwang *et al.* (1993) *Proc. Natl. Acad. Sci. USA*, 90:6340-6344; Yamada *et al.* (1994) *Human Gene Therapy,* 1:39-45; Leavitt *et al.* (1995) *Proc. Natl. Acad. Sci. USA,* 92:699-703; Leavitt *et al.* (1994) *Human Gene Therapy,* 5:1151-120; and Yamada *et al.* (1994) *Virology,* 205:121-126).

2. Inhibitors of StAR Polypeptide Activity

StAR activity can also be decreased by the addition of an inhibitor of the StAR polypeptide. This can be accomplished in any of a number of ways, including by

providing a dominant negative StAR polypeptide, *e.g.*, a form of StAR that itself has no activity and which, when present in the same cell as a functional StAR, reduces or eliminates the StAR activity of the functional StAR. Design of dominant negative forms is well known to those of skill and is described, *e.g.*, in Herskowitz (1987) *Nature*, 329(6136):219-22. Also, inactive polypeptide variants (muteins) can be used, *e.g.*, by screening for the ability to inhibit StAR activity. Methods of making muteins are well known to those of skill (*see, e.g.*, U.S. Patent Nos. 5,486,463, 5,422,260, 5,116,943, 4,752,585, 4,518,504). In addition, any small molecule, *e.g.*, any peptide, amino acid, nucleotide, lipid, carbohydrate, or any other organic or inorganic molecule can be screened for the ability to bind to or inhibit StAR activity, as described below.

a) Screening for StAR Inhibitors

In one embodiment, this invention provides methods of screening for agents that modulate and preferably downregulate StAR protein activity. Preferred "screening" methods of this invention involve (i) contacting an StAR-expressing cell (*e.g.*, a cell capable of expressing StAR) with a test agent; and (ii) detecting the level of StAR activity (*e.g.*, as described above), where a decreased level of StAR activity as compared to the level of StAR activity in a cell not contacted with the test agent indicates that the test agent inhibits or downregulates StAR.

Virtually any agent can be tested in such an assay. Such agents include, but are not limited to, natural or synthetic nucleic acids, natural or synthetic polypeptides, natural or synthetic lipids, natural or synthetic small organic molecules, and the like. In one preferred format, test agents are provided as members of a combinatorial library. In preferred embodiments, a collection of small molecule inhibitors are tested for StAR inhibiting ability. A "small molecule inhibitor" of StAR is any molecule, *e.g.*, a carbohydrate, nucleotide, amino acid, oligonucleotide, oligopeptide, lipid, inorganic compound, *etc.* that inhibits StAR protein activity. Such molecules can inhibit StAR protein activity by any of a number of mechanisms, *e.g.*, by binding to an StAR protein and competitively inhibiting its interaction with mitochondria, with cholesterol, or with other proteins. Preferably, such "small molecule inhibitors" are smaller than about 10 kD. More preferably, such inhibitors are smaller than about 5, 2, or 1 kD or less.

As discussed above, test agents can be screened based on any of a number of factors, including, but not limited to, a level of StAR polynucleotide, *e.g.*, mRNA or gDNA,

the level of StAR polypeptide, the degree of binding of a compound to a StAR polynucleotide or polypeptide, StAR intracellular localization, or any functional properties of StAR protein, such as the ability of StAR activity to effect cholesterol translocation into the mitochondria and resulting steroid hormone synthesis. Such direct and indirect measures of StAR activity *in vivo* can readily be used to detect and screen for molecules that modulate StAR activity.

(i) Combinatorial Libraries

In certain embodiments, combinatorial libraries of potential StAR modulators will be screened for an ability to bind to a StAR polypeptide or to modulate StAR activity. Conventionally, new chemical entities with useful properties are generated by identifying a chemical compound (called a "lead compound") with some desirable property or activity, *e.g.*, StAR inhibiting activity, creating variants of the lead compound, and evaluating the property and activity of those variant compounds. However, the current trend is to shorten the time scale for all aspects of drug discovery. Because of the ability to test large numbers quickly and efficiently, high throughput screening (HTS) methods are replacing conventional lead compound identification methods.

In one preferred embodiment, high throughput screening methods involve providing a library containing a large number of potential therapeutic compounds (candidate compounds). Such "combinatorial chemical libraries" are then screened in one or more assays to identify those library members (particular chemical species or subclasses) that display a desired characteristic activity. The compounds thus identified can serve as conventional "lead compounds" or can themselves be used as potential or actual therapeutics.

A combinatorial chemical library is a collection of diverse chemical compounds generated by either chemical synthesis or biological synthesis by combining a number of chemical "building blocks" such as reagents. For example, a linear combinatorial chemical library, such as a polypeptide (*e.g.*, mutein) library, is formed by combining a set of chemical building blocks called amino acids in every possible way for a given compound length (*i.e.*, the number of amino acids in a polypeptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks (Gallop *et al.* (1994) *J. Med. Chem.*, 37(9):1233-1251).

Preparation and screening of combinatorial chemical libraries is well known to those of skill in the art. Such combinatorial chemical libraries include, but are not limited to, peptide libraries (*see, e.g.,* U.S. Patent No. 5,010,175, Furka (1991) *Int. J. Pept. Prot. Res.*, 37:487-493, Houghton *et al.* (1991) *Nature*, 354:84-88), peptoids (PCT Publication No WO 91/19735, 26 Dec. 1991), encoded peptides (PCT Publication WO 93/20242, 14 Oct. 1993), random bio-oligomers (PCT Publication WO 92/00091, 9 Jan. 1992), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs *et al.*, (1993) *Proc. Nat. Acad. Sci. USA*, 90:6909-6913), vinylogous polypeptides (Hagihara *et al.* (1992) *J. Amer. Chem. Soc.*, 114:6568), nonpeptidal peptidomimetics with a Beta-D-Glucose scaffolding (Hirschmann *et al.*, (1992) *J. Amer. Chem. Soc.* 114:9217-9218), analogous organic syntheses of small compound libraries (Chen *et al.* (1994) *J. Amer. Chem. Soc.*, 116:2661), oligocarbamates (Cho, et al., (1993) *Science*, 261:1303), and/or peptidyl phosphonates (Campbell *et al.*, (1994) *J. Org. Chem.*, 59:658). *See, generally,* Gordon *et al.*, (1994) *J. Med. Chem.*, 37:1385, nucleic acid libraries (*see, e.g.,* Strategene, Corp.), peptide nucleic acid libraries (*see, e.g.,* U.S. Patent 5,539,083), antibody libraries (*see, e.g.,* Vaughn *et al.* (1996) *Nature Biotechnology*, 14(3):309-314), and PCT/US96/10287), carbohydrate libraries (*see, e.g.,* Liang *et al.*, (1996) *Science*, 274:1520-1522, and U.S. Patent No. 5,593,853), and small organic molecule libraries (*see, e.g.,* benzodiazepines, Baum (1993) C&EN, Jan 18, page 33; isoprenoids, U.S. Patent No. 5,569,588; thiazolidinones and metathiazanones, U.S. Patent No. 5,549,974; pyrrolidines, U.S. Patent Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Patent No. 5,506,337; benzodiazepines, U.S. Patent No. 5,288,514; and the like).

Devices for the preparation of combinatorial libraries are commercially available (*see, e.g.,* 357 MPS, 390 MPS, Advanced Chem Tech, Louisville KY, Symphony, Rainin, Woburn, MA, 433A Applied Biosystems, Foster City, CA, 9050 Plus, Millipore, Bedford, MA).

A number of well known robotic systems have also been developed for solution phase chemistries. These systems include automated workstations like the automated synthesis apparatus developed by Takeda Chemical Industries, LTD. (Osaka, Japan) and many robotic systems utilizing robotic arms (Zymate II, Zymark Corporation, Hopkinton, Mass.; Orca, Hewlett-Packard, Palo Alto, Calif.), which mimic the manual

synthetic operations performed by a chemist. Any of the above devices are suitable for use with the present invention. The nature and implementation of modifications to these devices (if any) so that they can operate as discussed herein will be apparent to persons skilled in the relevant art. In addition, numerous combinatorial libraries are themselves commercially available (*see, e.g.*, ComGenex, Princeton, N.J., Asinex, Moscow, Ru, Tripos, Inc., St. Louis, MO, ChemStar, Ltd, Moscow, RU, 3D Pharmaceuticals, Exton, PA, Martek Biosciences, Columbia, MD, *etc.*).

(ii) High Throughput Screening

Any of the assays to identify compounds capable of modulating StAR levels described herein are amenable to high throughput screening. Preferred assays thus detect enhancement or inhibition of StAR gene transcription, inhibition or enhancement of StAR polypeptide expression, and inhibition or enhancement of StAR polypeptide activity.

High throughput assays for the presence, absence, quantification, or other properties of particular nucleic acids or protein products are well known to those of skill in the art. Similarly, binding assays and reporter gene assays are similarly well known. Thus, for example, U.S. Patent No. 5,559,410 discloses high throughput screening methods for proteins, U.S. Patent No. 5,585,639 discloses high throughput screening methods for nucleic acid binding (*i.e.*, in arrays), while U.S. Patent Nos. 5,576,220 and 5,541,061 disclose high throughput methods of screening for ligand/antibody binding.

In addition, high throughput screening systems are commercially available (*see, e.g.*, Zymark Corp., Hopkinton, MA; Air Technical Industries, Mentor, OH; Beckman Instruments, Inc. Fullerton, CA; Precision Systems, Inc., Natick, MA, *etc.*). These systems typically automate entire procedures, including all sample and reagent pipetting, liquid dispensing, timed incubations, and final readings of the microplate in detector(s) appropriate for the assay. These configurable systems provide high throughput and rapid start up as well as a high degree of flexibility and customization. The manufacturers of such systems provide detailed protocols for various high throughput systems. Thus, for example, Zymark Corp. provides technical bulletins describing screening systems for detecting the modulation of gene transcription, ligand binding, and the like.

B. ADMINISTRATION OF STAR-INHIBITING COMPOUNDS

In numerous embodiments of the present invention, an StAR inhibiting compounds, *i.e.*, a compound that reduces levels of StAR mRNA, polypeptide and/or protein activity, will be administered to an animal. Such compounds can be administered by a variety of methods including, but not limited to, parenteral, topical, oral, or local administration, such as by aerosol or transdermally, for prophylactic and/or therapeutic treatment. The pharmaceutical compositions can be administered in a variety of unit dosage forms depending upon the method of administration. For example, unit dosage forms suitable for oral administration include, but are not limited to, powder, tablets, pills, capsules and lozenges. It is recognized that the StAR modulators (*e.g.*, antibodies, antisense constructs, ribozymes, small organic molecules, *etc.*) when administered orally, must be protected from digestion. This is typically accomplished either by complexing the molecule(s) with a composition to render it resistant to acidic and enzymatic hydrolysis, or by packaging the molecule(s) in an appropriately resistant carrier, such as a liposome. Means of protecting agents from digestion are well known in the art.

The compositions for administration will commonly comprise a StAR modulator dissolved in a pharmaceutically acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers can be used, *e.g.*, buffered saline and the like. These solutions are sterile and generally free of undesirable matter. These compositions may be sterilized by conventional, well known sterilization techniques. The compositions may contain pharmaceutically acceptable auxiliary substances as required to approximate physiological conditions such as pH adjusting and buffering agents, toxicity adjusting agents and the like, for example, sodium acetate, sodium chloride, potassium chloride, calcium chloride, sodium lactate and the like. The concentration of active agent in these formulations can vary widely, and will be selected primarily based on fluid volumes, viscosities, body weight and the like in accordance with the particular mode of administration selected and the patient's needs.

Thus, a typical pharmaceutical composition for intravenous administration would be about 0.1 to 10 mg per patient per day. Dosages from 0.1 up to about 100 mg per patient per day may be used, particularly when the drug is administered to a secluded site and not into the blood stream, such as into a body cavity or into a lumen of an organ. Substantially higher dosages are possible in topical administration. Actual methods for preparing parenterally administrable compositions will be known or apparent to those skilled

in the art and are described in more detail in such publications as *Remington's Pharmaceutical Science*, 15th ed., Mack Publishing Company, Easton, Pennsylvania (1980).

The compositions containing modulators of StAR can be administered for therapeutic or prophylactic treatments. In therapeutic applications, compositions are administered to a patient suffering from a disease (*e.g.*, a cancer) in an amount sufficient to cure or at least partially arrest the disease and its complications. An amount adequate to accomplish this is defined as a "therapeutically effective dose." Amounts effective for this use will depend upon the severity of the disease and the general state of the patient's health. Single or multiple administrations of the compositions may be administered depending on the dosage and frequency as required and tolerated by the patient. In any event, the composition should provide a sufficient quantity of the agents of this invention to effectively treat the patient. An amount of a StAR modulator that is capable of preventing or slowing the development of cancer in a mammal is referred to as a "prophylactically effective dose." The particular dose required for a prophylactic treatment will depend upon the medical condition and history of the mammal, the particular cancer being prevented, as well as other factors such as age, weight, gender, etc. Such prophylactic treatments may be used, *e.g.*, in a mammal who has previously had cancer to prevent a recurrence of the cancer, or in a mammal who is suspected of having a significant likelihood of developing cancer.

It will be appreciated that any of the present StAR-inhibiting compounds can be administered alone or in combination with additional StAR-inhibiting compounds or with any other therapeutic agent, *e.g.*, other anti-cancer agents or treatments.

1. Introducing Nucleic Acids into Cells

In numerous embodiments, one or more nucleic acids, *e.g.*, StAR polynucleotides, such as antisense polynucleotides or ribozymes, will be introduced into cells, *in vitro* or *in vivo*. The present invention provides methods, reagents, vectors, and cells useful for expression of StAR and other polypeptides and nucleic acids using *in vitro* (cell-free), *ex vivo* or *in vivo* (cell or organism-based) recombinant expression systems.

The particular procedure used to introduce the nucleic acids into a host cell for expression of a protein or nucleic acid is application specific. Any of the well known procedures for introducing foreign nucleotide sequences into host cells may be used. These include the use of calcium phosphate transfection, spheroplasts, electroporation, liposomes, microinjection, plasma vectors, viral vectors and any of the other well known methods for

introducing cloned genomic DNA, cDNA, synthetic DNA or other foreign genetic material into a host cell (*see, e.g.*; Berger and Kimmel, *Guide to Molecular Cloning Techniques, Methods in Enzymology* volume 152 Academic Press, Inc., San Diego, CA (Berger), F.M. Ausubel *et al.*, eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 1999), and Sambrook *et al.*, *Molecular Cloning - A Laboratory Manual* (2nd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989.

a) Vectors

In numerous embodiments of this invention, nucleic acids encoding StAR polypeptides, or inhibitors thereof, will be inserted into vectors using standard molecular biological techniques. Vectors may be used at multiple stages of the practice of the invention, *e.g.*, for subcloning nucleic acids encoding StAR polypeptides or StAR inhibitors, *e.g.*, StAR ribozymes or antisense sequences, or for subcloning additional elements used to control protein or mRNA expression, vector selectability, *etc.* Vectors may also be used to maintain or amplify the nucleic acids, for example, by inserting the vector into prokaryotic or eukaryotic cells and growing the cells in culture. In addition, vectors may be used to introduce and express StAR nucleic acids, or StAR-inhibiting nucleic acids, *e.g.*, StAR ribozymes or antisense sequences, into cells for therapeutic or experimental purposes.

A variety of commercially or commonly available vectors and vector nucleic acids can be converted into a vector of the invention by cloning a polynucleotide of this invention into the commercially or commonly available vector. A variety of common vectors suitable for this purpose are well known in the art. For cloning in bacteria, common vectors include pBR322-derived vectors such as pBLUESCRIPT™, and bacteriophage derived vectors. In yeast, vectors include Yeast Integrating plasmids (*e.g.*, YIp5) and Yeast Replicating plasmids (the YRp series plasmids) and pGPD-2. Expression in mammalian cells can be achieved using a variety of commonly available plasmids, including pSV2, pBC12BI, and p91023, as well as lytic virus vectors (*e.g.*, vaccinia virus, adeno virus, and baculovirus), episomal virus vectors (*e.g.*, bovine papillomavirus), and retroviral vectors (*e.g.*, murine retroviruses).

Typically, a nucleic acid subsequence encoding an StAR polypeptide is placed under the control of a promoter. A nucleic acid is "operably linked" to a promoter when it is placed into a functional relationship with the promoter. For instance, a promoter

or enhancer is operably linked to a coding sequence if it increases or otherwise regulates the transcription of the coding sequence. Similarly, a "recombinant expression cassette" or simply an "expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with nucleic acid elements that are capable of effecting expression of a structural gene in hosts compatible with such sequences. Expression cassettes include promoters and, optionally, introns, polyadenylation signals, and transcription termination signals. Additional factors necessary or helpful in effecting expression may also be used as described herein. For example, an expression cassette can also include nucleotide sequences that encode a signal sequence that directs secretion of an expressed protein from the host cell. Transcription termination signals, enhancers, and other nucleic acid sequences that influence gene expression, can also be included in an expression cassette.

An extremely wide variety of promoters are well known, and can be used in the vectors of the invention, depending on the particular application. Ordinarily, the promoter selected depends upon the cell in which the promoter is to be active. Other expression control sequences such as ribosome binding sites, transcription termination sites and the like are also optionally included. For *E. coli*, example control sequences include the T7, trp, or lambda promoters, a ribosome binding site and preferably a transcription termination signal. For eukaryotic cells, the control sequences typically include a promoter which optionally includes an enhancer derived from immunoglobulin genes, SV40, cytomegalovirus, a retrovirus (*e.g.*, an LTR based promoter) etc., and a polyadenylation sequence, and may include splice donor and acceptor sequences.

VI. Cancers

The present methods can be used to diagnose and treat any of a number of types of cancers. In preferred embodiments, epithelial cancers will be diagnosed and/or treated, *e.g.*, breast cancer. Other epithelial cancers include, *e.g.*, ovarian, colorectal, kidney, stomach, bladder, and lung cancers. A cancer at any stage of progression can be detected, such as primary, metastatic, and recurrent cancers. Information regarding numerous types of cancer can be found, *e.g.*, from the American Cancer Society (www3.cancer.org), or from, *e.g.*, Wilson *et al.* (1991) Harrison's Principles of Internal Medicine, 12th Edition, McGraw-Hill, Inc.

VII. Kits for Use in Diagnostic and/or Prognostic Applications.

For use in diagnostic, research, and therapeutic applications suggested above, kits are also provided by the invention. In the diagnostic and research applications such kits may include any or all of the following: assay reagents, buffers, StAR specific nucleic acids or antibodies, hybridization probes and/or primers, antisense polynucleotides, ribozymes, dominant negative StAR polypeptides or polynucleotides, small molecules inhibitors of StAR, *etc*. A therapeutic product may include sterile saline or another pharmaceutically acceptable emulsion and suspension base.

In addition, the kits may include instructional materials containing directions (*i.e.*, protocols) for the practice of the methods of this invention. While the instructional materials typically comprise written or printed materials they are not limited to such. Any medium capable of storing such instructions and communicating them to an end user is contemplated by this invention. Such media include, but are not limited to electronic storage media (*e.g.*, magnetic discs, tapes, cartridges, chips), optical media (*e.g.*, CD ROM), and the like. Such media may include addresses to internet sites that provide such instructional materials.

VIII. Transgenic Animals

Transgenic and chimeric non-human mammals and methods for generating them are described below. The mammals are useful, *inter alia*, for testing the function of StAR *in vivo*, to generate models for the study of diseases and conditions associated with StAR (*e.g.*, cancer), and for the development of potential treatments for diseases and conditions associated with StAR (*e.g.*, cancer).

Transgenic and chimeric non-human mammals are generated that contain cells that lack at least one functional endogenous allele for StAR. A "chimeric animal" includes some cells that lack the functional StAR gene of interest and other cells that do not have the inactivated gene. A "transgenic animal," in contrast, is made up of cells that have all incorporated the specific modification which renders the StAR gene inactive or otherwise altered. While a transgenic animal is typically always capable of transmitting the mutant StAR gene to its progeny, the ability of a chimeric animal to transmit the mutation depends upon whether the inactivated gene is present in the animal's germ cells. The modifications that inactivate or otherwise alter the StAR gene can include, for example, insertions,

deletions, or substitutions of one or more nucleotides. The modifications can interfere with transcription of the gene itself, with translation and/or stability of the resulting mRNA, or can cause the gene to encode an inactive or otherwise altered StAR polypeptide.

The methods of the present invention are useful for producing transgenic and chimeric animals of most vertebrate species. Such species include, but are not limited to, nonhuman mammals, including rodents such as mice and rats, rabbits, ovines such as sheep and goats, porcines such as pigs, and bovines such as cattle and buffalo. Methods of obtaining transgenic animals are described in, for example, Puhler, A., Ed., *Genetic Engineering of Animals*, VCH Publ., 1993; Murphy and Carter, Eds., *Transgenesis Techniques : Principles and Protocols* (*Methods in Molecular Biology*, Vol. 18), 1993; and Pinkert, CA, Ed., *Transgenic Animal Technology : A Laboratory Handbook*, Academic Press, 1994.

In preferred embodiments, transgenic mice will be produced as described in Thomas *et al.*, (1999) *Immunol.*, 163:978-84; Kanakaraj *et al.* (1998) *J. Exp. Med.*, 187:2073-9; or Yeh *et al.*, (1997) *Immunity*, 7:715-725.

Typically, a modified StAR gene is introduced, *e.g.*, by homologous recombination, into embryonic stem cells (ES), which are obtained from preimplantation embryos and cultured *in vitro*. *See, e.g.*, Hooper, ML, *Embryonal Stem Cells : Introducing Planned Changes into the Animal Germline* (Modern Genetics, v. 1), Int'l. Pub. Distrib., Inc., 1993; Bradley *et al.* (1984) *Nature*, 309, 255-258. Subsequently, the transformed ES cell is combined with a blastocyst from a non-human animal, *e.g.*, a mouse. The ES cells colonize the embryo and in some embryos form the germ line of the resulting chimeric animal. *See*, Jaenisch (1988) *Science*, 240:1468-1474. Alternatively, ES cells or somatic cells that can reconstitute an organism ("somatic repopulating cells") can be used as a source of nuclei for transplantation into an enucleated fertilized oocyte giving rise to a transgenic mammal. *See, e.g.*, Wilmut *et al.* (1997) *Nature* 385:810-813.

Other methods for obtaining a transgenic or chimeric animal having a mutant StAR gene in its genome is to contact fertilized oocytes with a vector that includes a polynucleotide that encodes a modified, *e.g.*, inactive, StAR polypeptide. In some animals, such as mice, fertilization is typically performed *in vivo* and fertilized ova are surgically removed. In other animals, particularly bovines, it is preferably to remove ova from live or slaughterhouse animals and fertilize the ova *in vitro*. *See*, DeBoer *et al.*, WO 91/08216. *In*

vitro fertilization permits the modifications to be introduced into substantially synchronous cells.

Fertilized oocytes are typically cultured *in vitro* until a pre-implantation embryo is obtained containing about 16-150 cells. The 16-32 cell stage of an embryo is described as a morula, whereas pre-implantation embryos containing more than 32 cells are termed blastocysts. These embryos show the development of a blastocoel cavity, typically at the 64 cell stage. The presence of a desired StAR mutation in the cells of the embryo can be detected by methods known to those of skill in the art, *e.g.*, Southern blotting, PCR, DNA sequencing, or other standard methods. Methods for culturing fertilized oocytes to the pre-implantation stage are described, *e.g.*, by Gordon *et al.* (1984) *Methods Enzymol.*, 101:414; Hogan *et al. Manipulation of the Mouse Embryo: A Laboratory Manual*, C.S.H.L. N.Y. (1986) (mouse embryo); Hammer *et al.* (1985) *Nature*, 315:680 (rabbit and porcine embryos); Gandolfi *et al.* (1987) *J. Reprod. Fert.* 81:23-28; Rexroad *et al.* (1988) *J. Anim. Sci.*, 66:947-953 (ovine embryos) and Eyestone *et al.* (1989) *J. Reprod. Fert.*, 85:715-720; Camous *et al.* (1984) *J. Reprod. Fert.*, 72:779-785; and Heyman *et al.* (1987) *Theriogenology*, 27:5968 (bovine embryos). Pre-implantation embryos may also be stored frozen for a period pending implantation.

Pre-implantation embryos are transferred to an appropriate female resulting in the birth of a transgenic or chimeric animal, depending upon the stage of development when the transgene is integrated. Chimeric mammals can be bred to form true germline transgenic animals. Chimeric mice and germline transgenic mice can also be ordered from commercial sources (*e.g.*, Deltagen, San Carlos, CA).

Other methods for introducing mutations into mammalian cells or animals include recombinase systems, which can be employed to delete all or a portion of a locus of interest. Examples of recombinase systems include, the cre/lox system of bacteriophage P1 (*see, e.g.*, Gu *et al.* (1994) *Science*, 265:103-106; Terry *et al.* (1997) *Transgenic Res.*, 6:349-356) and the FLP/FRT site specific integration system (*see, e.g.*, Dymecki (1996) *Proc. Natl. Acad. Sci. USA*, 93:6191-6196). In these systems, sites recognized by the particular recombinase are typically introduced into the genome at a position flanking the portion of the gene that is to be deleted. Introduction of the recombinase into the cells then catalyzes recombination which deletes from the genome the polynucleotide sequence that is flanked by the recombination sites. If desired, one can obtain animals in which only certain cell

types lack the StAR gene of interest, *e.g.*, by using a tissue specific promoter to drive the expression of the recombinase. *See, e.g.*, Tsien *et al.* (1996) *Cell,* 87:1317-26; Brocard *et al.* (1996) *Proc. Natl. Acad. Sci. USA,* 93:10887-10890; Wang *et al.* (1996) *Proc. Natl. Acad. Sci. USA,* 93:3932-6; Meyers *et al.* (1998) *Nat. Genet.*, 18:136-41).

The presence of any mutation in a StAR gene in a cell or animal can be detected using any method described herein, *e.g.*, Southern blot, PCR, or DNA sequencing. *See, e.g.*, Ausubel *et al.*, *supra*.

EXAMPLES

Discovery of novel amplified genes in breast cancer

By representational difference analysis (RDA) (Lisitsyn *et al.*, *supra*) of a human breast cancer biopsy, an RDA probe was discovered that maps to 8p11 and that detects amplification in several breast cancer samples. Amplification of 8p11 has been detected previously and is associated with amplification of *FGFR1*, although not all tumors amplified for *FGFR1* overexpress the gene (Dib *et al.*, (1995) *Oncogene,* 10:995-1001; Ugolini *et al.*, (1999) *Oncogene,* 18:1903-10). Several results strongly indicate that *FGFR1* is not the only oncogene that is driving amplification of this region. First, the cell line BT483 is not amplified for *FGFR1,* but is amplified for the RDA probe. Second, one tumor (89-249) that was analyzed by quantitative RT-PCR does not overexpress the *FGFR1* gene, indicating that *FGFR1* is not the oncogene responsible for amplification of the region in this tumor. Third, the physical map of this region reveals that *FGFR1* is approximately 0.6 Mb away from the RDA probe, and there are sequences between the RDA probe and *FGFR1* that are not amplified or less amplified than the RDA probe and *FGFR1* (Figure 1). Thus, the region amplified near the RDA probe contains a separate oncogene.

Significant overexpression of *ASH2, StAR, CaSM, and SODD* in amplified tumors

Based on more detailed analysis of primary tumors shown in Figure 1, and analysis of 9 additional amplified primary tumors, the smallest area of common overlap of amplification was determined to encompass four genes: *ASH2*, the human homolog of a Drosophila trithorax gene (Ikegawa *et al.*, (1999) *Cytogenet. Cell Genet.,* 84:167-72), *StAR*, which encodes the steroidogenic acute regulatory protein (Clark *et al.*, *supra*), *CaSM*, which encodes a protein with Sm-like motifs that are characteristic of small nuclear ribonucleoprotein particle proteins (Schweinfest *et al.*, (1997) *Cancer Res.*, 57:2961-5), and

SODD, which was recently discovered as a protein that binds to the TNF-receptor's death domain and that inhibits TNF-induced apoptosis (Jiang *et al.*, 1999 *Science*, 283:543-6; 283:1852).

All four of these genes are overexpressed in amplified tumors based on quantitative RT-PCR analysis using TaqMan probes (Table 1).

Table 1. TaqMan quantitative RT-PCR analysis of expression levels in human breast cancer biopsies of genes from the 8p11 amplified region

		Relative Expression Level			
Tumor or tissue	Amplification	ASH2	StAR	CaSM	SODD
Normal Mammary	-	1	1	1	1
88-688	+	12	17	24	11
89-249	+	5.1	6.2	11	5.0
96-76	+	7.0	3.5	11	4.4
88-523	+	12	26	18	5.9
96-201	+	10	9.0	19	4.6
96-194	-	0.8	1.6	1.3	0.9
96-360	-	1.4	1.4	2.1	0.9

TaqMan RT-PCR analysis of total RNA was performed according to protocols from Perkin Elmer, Foster City, CA. Relative expression level is the ratio of the value determined for the test gene relative to the value determined for the reference gene β-actin, after normalization to the ratio determined for normal human mammary gland (Clontech, Palo Alto, CA).

One method to discern which overexpressed gene is driving amplification of the region is to look for overexpression of the genes in the absence of gene amplification. For example, several important amplified oncogenes, such as MYC and CCDN1 (encoding Cyclin D1), are frequently overexpressed in the absence of gene amplification, presumably as a result of other primary genetic alterations. Overexpression in the absence of amplification argues that it is the overexpression of the gene that is important, as opposed to its physical location near a true amplified cancer gene that underlies overexpression.

Of the four genes, StAR stands outs as the most frequently and most highly overexpressed gene in nonamplified tumors based on quantitative RT-PCR analysis using TaqMan probes (Table 2).

Table 2. TaqMan quantitative RT-PCR analysis of expression levels in human breast cancer biopsies from nonamplified tumors of genes from the 8p11 amplified region

	Relative Expression Level			
Tumor or tissue	ASH2	StAR	CaSM	SODD
Normal Mammary	1	1	1	1
88-486	0.4	0.4	0.7	< 1
88-682	1.3	2.0	2.9	< 1
96-342	0.8	6.2	4.4	0.7
96-349	1.1	2.6	2.0	< 1
96-317	0.6	1.6	1.5	< 1
96-273	2.1	1.6	2.1	0.4
96-252	1.0	2.2	1.7	0.4
96-190	1.5	0.4	3.4	0.8
96-160	0.5	3.2	1.9	0.6
96-140	0.8	0.8	1.1	< 1
96-109	0.9	3.6	0.8	< 1
96-102	0.3	0.5	0.5	< 1
96-32	0.6	0.6	0.5	< 1
96-16	0.4	0.4	0.6	0.1
95-523	0.4	0.5	0.7	0.2
95-504	0.6	1.2	0.6	0.3
95-487	1.9	9.2	0.7	0.4
95-480	2.2	1.0	2.1	0.9
95-476	0.6	0.6	0.6	< 1
95-427	0.9	2.9	2.3	0.6
95-377	0.9	2.0	1.8	0.4
95-326	1.4	8.2	2.7	0.7
95-283	1.6	3.3	1.6	0.7
95-237	1.3	2.8	1.9	< 1
95-65	0.8	2.3	1.4	0.3
94-847	0.8	1.6	1.4	< 1
94-797	0.9	3.8	2.1	< 1
96-194	0.8	2	1	0.9
96-360	1	1	2	0.9

TaqMan RT-PCR analysis of total RNA was performed according to protocols from Perkin Elmer, Foster City, CA. Relative expression level is the ratio of the value determined for the test gene relative to the value determined for the reference gene β-actin, after normalization to the ratio determined for normal human mammary gland (Clontech, Palo Alto, CA).

The overexpression of StAR in human tumors is strongly associated with estrogen receptor and progesterone receptor status. Tumors that overexpress StAR are almost exclusively ER or PR positive, whereas tumors that express normal levels of StAR are predominantly ER and PR negative (Table 3).

Table 3. The expression of StAR in human breast tumors is associated with steroid receptor status

	Number of Tumors with Indicated Steroid Receptor Status				
Expression Levels of StAR	ER+/PR+	ER-/PR+	ER+/PR-	ER-/PR-	Not known
> 2.5	6	1	0	1	2
2 – 2.5	4	0	0	0	1
< 2	1	1	1	10	1

These findings and the known biochemical function of StAR in the adrenal gland leads to the hypothesis that StAR is the cancer gene driving amplification of the 8p11 region, and that its overexpression upregulates estrogen biosynthesis causing "autocrine" stimulation of breast cancer cell growth.

While the foregoing invention has been described in some detail for purposes of clarity and understanding, it will be clear to one skilled in the art from a reading of this disclosure that various changes in form and detail can be made without departing from the true scope of the invention. For example, all the techniques and apparatus described above may be used in various combinations. All publications and patent documents cited in this application are incorporated by reference in their entirety for all purposes to the same extent as if each individual publication or patent document were so individually denoted.

SEQ ID NO:1

Human StAR cDNA sequence

GGGACTCAGAGGCGAAGCTTGAGGGGCTCAGGAAGGACGAAGAACCACCCTTG
AGAGAAGAGGCAGCAGCAGCGGCGGCAGCAGCAGCGGCAGCGACCCCACCACT
GCCACATTTGCCAGGAAACAATGCTGCTAGCGACATTCAAGCTGTGCGCTGGGA
GCTCCTACAGACACATGCGCAACATGAAGGGGCTGAGGCAACAGGCTGTGATG
GCCATCAGCCAGGAGCTGAACCGGAGGGCCCTGGGGGGCCCCACCCCTAGCAC
GTGGATTAACCAGGTTCGGCGGCGGAGCTCTCTACTCGGTTCTCGGCTGGAAGA
GACTCTCTACAGTGACCAGGAGCTGGCCTATCTCCAGCAGGGGGAGGAGGCCAT
GCAGAAGGCCTTGGGCATCCTTAGCAACCAAGAGGGCTGGAAGAAGGAGAGTC
AGCAGGACAATGGGGACAAAGTGATGAGTAAAGTGGTCCCAGATGTGGGCAAG
GTGTTCCGGCTGGAGGTCGTGGTGGACCAGCCCATGGAGAGGCTCTATGAAGAG
CTCGTGGAGCGCATGGAAGCAATGGGGGAGTGGAACCCCAATGTCAAGGAGAT
CAAGGTCCTGCAGAAGATCGGAAAAGATACATTCATTACTCACGAGCTGGCTGC
CGAGGCAGCAGGAAACCTGGTGGGGCCCCGTGACTTTGTGAGCGTGCGCTGTGC
CAAGCGCCGAGGCTCCACCTGTGTGCTGGCTGGCATGGACACAGACTTCGGGAA
CATGCCTGAGCAGAAGGGTGTCATCAGGGCGGAGCACGGTCCCACTTGCATGGT
GCTTCACCCGTTGGCTGGAAGTCCCTCTAAGACCAAACTTACGTGGCTACTCAGC
ATCGACCTCAAGGGGTGGCTGCCCAAGAGCATCATCAACCAGGTCCTGTCCCAG
ACCCAGGTGGATTTTGCCAACCACCTGCGCAAGCGCCTGGAGTCCCACCCTGCC
TCTGAAGCCAGGTGTTGAAGACCAGCCTGCTGTTCCCAACTGTGCCCAGCTGCA
CTGGTACACACGCTCATCAGGAGAATCCCTACTGGAAGCCTGCAAGTCTAAGAT
CTCCATCTGGTGACAGTGGGATGGGTGGGGTTCGTGTTTAGAGTATGACACTAG
GATTCAGATTGGTGAAGTTTTTAGTACCAAGAAAACAGGGATGAGGCTCTTGGA
TTAAAAGGTAACTTCATTCACTGATTAGCTATGACATGAGGGTTCAGGCCCCTA
AAATAATTGTAAAACTTTTTTTCTGGGCCCTTATGTACCCACCTAAAACCATCTT
TAAAATGCTAGTGGCTGATATGGGTGTGGGGGATGCTAACCACAGGGCCTGAGA
AGTCTTGCTTTATGGGCTCAAGAATGCCATGCGCTGGCAGTACATGTGCACAAA
GCAGAATCTCAGAGGGTCTCCTGCAGCCCTCTGCTCCTCCCGGCCGCTGCACAG
CAACACCACAGAACAAGCAGCACCCCACAGTGGGTGCCTTCCAGAAATATAGTC
CAAGCTTTCTCTGTGGAAAAAGACAAAACTCATTAGTAGACATGTTTCCCTATTG
CTTTCATAGGCACCAGTCAGAATAAAGAATCATAATTCACACC

SEQ ID NO:2

Human StAR amino acid sequence

MLLATFKLCAGSSYRHMRNMKGLRQQAVMAISQELNRRALGGPTPSTWINQVRRR
SSLLGSRLEETLYSDQELAYLQQGEEAMQKALGILSNQEGWKKESQQDNGDKVMS
KVVPDVGKVFRLEVVVDQPMERLYEELVERMEAMGEWNPNVKEIKVLQKIGKDTF
ITHELAAEAAGNLVGPRDFVSVRCAKRRGSTCVLAGMDTDFGNMPEQKGVIRAEH
GPTCMVLHPLAGSPSKTKLTWLLSIDLKGWLPKSIINQVLSQTQVDFANHLRKRLES
HPASEARC"

WHAT IS CLAIMED IS:

1. A method of detecting cancer in a biological sample from a mammal, the method comprising detecting the presence or absence of a StAR polypeptide or polynucleotide in said biological sample, wherein a diagnostic presence of said StAR polypeptide or polynucleotide indicates the presence of cancer in said biological sample.

2. The method of claim 1, wherein the detection is performed under conditions that would not detect said StAR polypeptide or polynucleotide in a biological sample that is not cancerous, and wherein a diagnostic presence comprises detecting any StAR polypeptide or polynucleotide in said biological sample.

3. The method of claim 1, wherein said diagnostic presence comprises at least a 2-fold increase in the level of said StAR polypeptide or polynucleotide in said biological sample compared to a level expected in a biological sample from a control, cancer-free mammal.

4. The method of claim 1, wherein said diagnostic presence comprises at least a 5-fold increase in the level of said StAR polypeptide or polynucleotide in said biological sample compared to a level expected in a biological sample from a control, cancer-free mammal.

5. The method of claim 1, wherein said diagnostic presence comprises at least a 10-fold increase in the level of said StAR polypeptide or polynucleotide in said biological sample compared to a level expected in a biological sample from a control, cancer-free mammal.

6. The method of claim 1, wherein the method comprises detecting a StAR polypeptide by immunoassay.

7. The method of claim 6, wherein the immunoassay comprises immunohistochemistry.

8. The method of claim 1, wherein the method comprises detecting the presence or absence of StAR gene amplification.

9. The method of claim 1, wherein the method comprises detecting StAR mRNA.

10. The method of claim 1, wherein said cancer is an epithelial cancer.

11. The method of claim 10, wherein said epithelial cancer is breast cancer.

12. The method of claim 1, wherein said mammal is a human.

13. The method of claim 1, wherein said biological sample is selected from the group consisting of tissue biopsy, blood sample, and nipple discharge.

14. A method of detecting cancer in a biological sample from a mammal, the method comprising detecting the level of a StAR polypeptide or polynucleotide in said biological sample, wherein an increased level of said StAR polypeptide or polynucleotide in said biological sample compared to the level expected of a control sample from a normal, cancer-free mammal indicates the presence of cancer in said biological sample.

15. The method of claim 14, wherein said increased level comprises at least a 2-fold increase in the level of said StAR polypeptide or polynucleotide in said biological sample compared to a level expected in a biological sample from a control, cancer-free mammal.

16. The method of claim 14, wherein said increased level comprises at least a 5-fold increase in the level of said StAR polypeptide or polynucleotide in said biological sample compared to a level expected in a biological sample from a control, cancer-free mammal.

17. The method of claim 14, wherein said increased level comprises at least a 10-fold increase in the level of said StAR polypeptide or polynucleotide in the biological sample compared to a level expected in a biological sample from a control, cancer-free mammal.

18. The method of claim 14, wherein the method comprises detecting a StAR polypeptide by immunoassay.

19. The method of claim 14, wherein the method comprises detecting the presence or absence of StAR gene amplification.

20. The method of claim 14, wherein the method comprises detecting StAR mRNA.

21. The method of claim 14, wherein said cancer is an epithelial cancer.

22. The method of claim 21, wherein said epithelial cancer is breast cancer.

23. The method of claim 14, wherein said biological sample is selected from the group consisting of tissue biopsy, blood sample, and nipple discharge.

24. The method of claim 14, wherein said mammal is a human.

25. A method of monitoring the efficacy of a cancer treatment, the method comprising detecting the level of a StAR polypeptide or polynucleotide in a biological sample from a mammal undergoing treatment for cancer, wherein a reduced level of said StAR polypeptide or polynucleotide in said biological sample compared to the level in a biological sample from the mammal prior to, or earlier in, the treatment is indicative of efficacious treatment.

26. The method of claim 25, wherein said cancer is an epithelial cancer.

27. The method of claim 26, wherein said epithelial cancer is breast cancer.

28. The method of claim 25, wherein said mammal is a human.

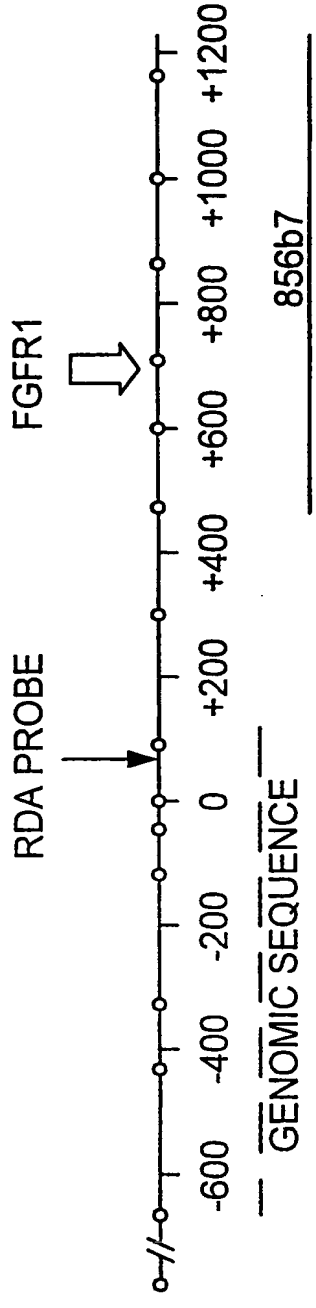
29. A method of decreasing the proliferation of a cell with an elevated level of StAR, said method comprising reducing the level of StAR activity in said cell using an inhibitor of StAR.

30. The method of claim 29, wherein said cell is a cancer cell.

31. The method of claim 30, wherein said cancer cell is an epithelial cancer cell.

32. The method of claim 31, wherein said epithelial cancer cell is a breast cancer cell.

33. The method of claim 29, wherein said inhibitor produces a decrease in the production of steroid hormones by the cell.



FOLD AMPLIFICATION:																			
87-425	2	1		4	4		9	7	12	5	6	2	3	6	6	2	2		
88-523	4	4		2	3		9	9	7	5	9	5	4	10	5	4	1		
88-688	2	10		3	7		10	9	11	8	6	7	8	8	10	8	9		
89-249	1	4		1	2		4	3	5	3	3	5	3	3	3	1	1		
96-201	2	2		1	2		3	5	5	4	2	1	3	3	1	2	1		
96-76	1	5		1	3		5	7	7	4	3	3	5	8	7	3	2		

FIG. 1.

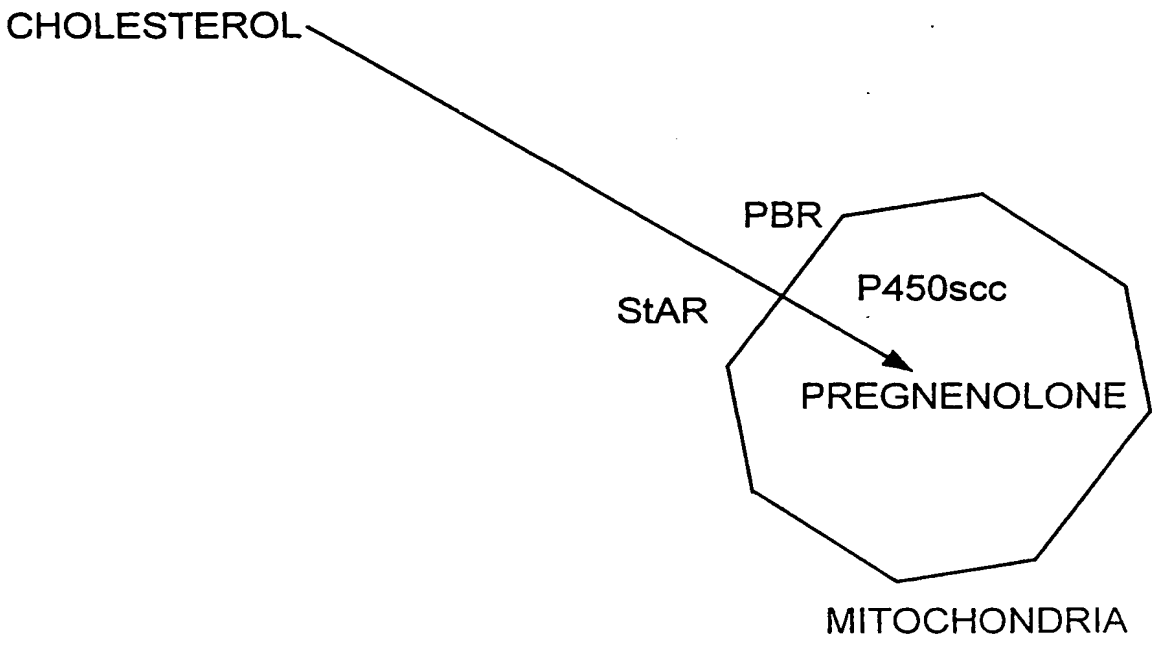


FIG. 2.